THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Volume 4 *No.* 16 *December* 1953

OXFORD

AT THE CLARENDON PRESS

Price 12*s.* 6*d.* *net*

PRINTED IN GREAT BRITAIN BY CHARLES BATEY AT THE UNIVERSITY PRESS, OXFORD

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Edited by T. W. CHAUNDY, U. S. HASLAM-JONES, J. H. C. THOMPSON

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SECOND SERIES) is published at 12*s.* 6*d.* net for a single number with an annual subscription (for four numbers) of 44*s.* 6*d.* post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford'. They should be in a condition fit for sending to a referee and in form and notation convenient for printing. The Editors as a rule will not wish to accept material that they cannot see their way to publish within a twelvemonth.

While every care is taken of manuscripts submitted for publication, the Publisher and the Editors cannot hold themselves responsible for any loss or damage. Authors are advised to retain a copy of anything they may send for publication. Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints.

Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the publisher:—

GEOFFREY CUMBERLEGE

OXFORD UNIVERSITY PRESS

AMEN HOUSE, LONDON, E.C. 4

THE INTERNATIONAL CONGRESS OF MATHEMATICIANS 1954

This Congress will be held in Amsterdam from Thursday, 2 September, to Thursday, 9 September (inclusive). Details of the plans for the Scientific Programme and the Entertainments and Excursions and of booking arrangements for Hotel Accommodation can be obtained from the Secretary. Correspondence should be addressed to

THE SECRETARIAT OF THE INTERNATIONAL
CONGRESS OF MATHEMATICIANS 1954

2e BOERHAAVESTRAAT 49

AMSTERDAM

THE NETHERLANDS

FIELDS OF PARALLEL PLANES IN AFFINELY CONNECTED SPACES

By Y. C. WONG (*Hong Kong*)

[Received 19 November 1952]

1. Introduction

STUDY of parallel partially null planes† in a Riemannian n-space V_n was initiated by Walker (**8**). Since then a number of papers on this subject have been published by Ruse, Patterson, and Walker himself (**4**, **5**, **6**, **9**, **10**). It should be mentioned, however, that parallel vectors, null or non-null, were considered much earlier by Eisenhart (**2**), who was also the first to use results from the theory of complete systems of partial differential equations‡ in studying problems dealing with parallel vectors.

In this paper, I give a few results on parallel planes in an affinely connected n-space A_n (with symmetric connexion), using largely Walker's notation and terminology for the case of V_n. The first half of this paper is devoted to the study of a certain 'double' tensor associated with each parallel contravariant plane in A_n (§ 3). This tensor is likely to play an important part in the theory of parallel planes, but I have so far obtained only some of its elementary properties, which are given in §§ 4–6.

The second half of this paper deals with a certain privileged coordinate system in an A_n admitting two parallel contravariant planes (§ 7). As applications, I first give a simpler derivation of Walker's canonical form for the metric of a V_n admitting a parallel partially null plane (§ 9), and then obtain a canonical form for the metric of an even-dimensional V_n admitting two non-intersecting parallel null $\frac{1}{2}n$-planes (§ 10). The latter V_n appears to deserve further study as its metric in canonical form happens to be the well-known Kähler metric in real variables.

Throughout this paper, the tensor indices i, j, k, l have the range $1,..., n$. Ranges of the scalar indices α, ξ, etc. which are used to number various vectors or tensors will be indicated at appropriate places.

2. Generalities

The definition and some of the basic properties of parallel planes in V_n, first given by Walker in (**8**), can be carried over to the case of A_n. I now state what I shall need in this paper.

† As usual, fields of parallel planes (or vectors) will be simply called parallel planes (or vectors).

‡ For a detailed account of the theory of complete systems, see Schouten and Kulk (**7**), where the lemma proved by Walker in (**9**) § 3 can also be found.

An r-plane at a point P of A_n is an r-dimensional (algebraic) vector space whose elements are vectors at P. The r-plane is said to be *contravariant* or *covariant* according as its vectors are all contravariant or all covariant. Since we shall concern ourselves mainly with contravariant r-planes, such planes will be simply called 'r-planes'.

An r-plane in A_n (i.e. a field of r-planes in A_n) is an r-plane uniquely and analytically defined over A_n; thus it may be determined by a *basis* consisting of r (linearly) independent vectors whose components λ^i_α are analytic functions of position. Such a basis will be denoted by $\{\lambda^i_\alpha\}$. In this and the next few sections, $\alpha, \beta, \gamma, \epsilon = 1,..., r$.

An r-plane Π in A_n is said to be *parallel* if, for any two points P and Q, a vector in $\Pi(P)$ is displaced into a vector in $\Pi(Q)$ by parallel transport along any curve from P to Q. It follows at once from this definition that, if Π, Σ are two parallel planes, the intersection $\Pi \cap \Sigma$ and the union $\Pi \cup \Sigma$ are both parallel planes.

A necessary and sufficient condition for an r-plane to be parallel is that one, and therefore every one, of its bases $\{\lambda^i_\alpha\}$ satisfies recurrent relations of the form [Walker (**8**), Theorem 3.3]

$$\lambda^i_{\alpha/k} = A^\beta_{\alpha k}\lambda^i_\beta, \tag{2.1}$$

where the solidus denotes covariant differentiation with respect to the (symmetric) connexion Γ^i_{jk} of A_n, and $A^\beta_{\alpha k}$ are covariant vectors. An important consequence of (2.1) is that, if $\{\lambda^i_\alpha\}$ is a basis of a parallel r-plane, the system of partial differential equations

$$\lambda^i_\alpha f_{.i} = 0 \quad \left(f_{.i} = \frac{\partial f}{\partial x^i}\right) \tag{2.2}$$

is complete, and therefore admits $n-r$ (functionally) independent solutions [Walker (**9**) 71]. It is obvious that replacing a basis by another corresponds to replacing the complete system by an equivalent system. I shall refer to any such complete system as *a complete system associated with the parallel r-plane*.

Similarly, a covariant $(n-r)$-plane is parallel if and only if it has a basis $\{\mu^\rho_j\}$ $(\rho, \sigma = r+1,..., n)$ satisfying recurrent relations of the form

$$\mu^\rho_{j/k} = B^\rho_{\sigma k}\mu^\sigma_j. \tag{2.3}$$

But, unlike the case of a parallel contravariant plane, there is no complete system of partial differential equations similar to (2.2) associated with a parallel covariant plane.

It can easily be verified that, if the vectors μ_j^ρ satisfy (2.3), any complete set of r independent solutions λ_α^i of the equations

$$\mu_i^\rho \lambda^i = 0$$

satisfies relations of the form (2.1). In this way, to each parallel covariant $(n-r)$-plane, there corresponds a unique parallel (contravariant) r-plane. Consequently, parallel covariant planes require no separate investigations; their properties can be inferred from those of parallel (contravariant) planes.

3. The 'double' tensor $A^\beta_{\alpha kl}$

Let Π be a parallel r-plane in A_n, and $\{\lambda_\alpha^i\}$ a basis of Π. Then

$$\lambda^i_{\alpha/k} = A^\beta_{\alpha k}\lambda^i_\beta, \tag{3.1}$$

where $A^\beta_{\alpha k}$ are covariant vectors.

Consider the change of basis

$$\lambda^{*i}_\alpha = c^\beta_\alpha \lambda^i_\beta, \tag{3.2}$$

or

$$\lambda^i_\alpha = \bar{c}^\beta_\alpha \lambda^{*i}_\beta, \tag{3.2$'$}$$

where $\det(c^\beta_\alpha) \neq 0$ and $\bar{c}^\beta_\alpha c^\gamma_\beta = \delta^\gamma_\alpha$. Differentiating (3.2) covariantly with respect to x^k and then using (3.1), we have

$$\lambda^{*i}_{\alpha/k} = A^{*\epsilon}_{\alpha k}\lambda^{*i}_\epsilon,$$

where

$$A^{*\epsilon}_{\alpha k} = (c^\gamma_{\alpha/k} + c^\beta_\alpha A^\gamma_{\beta k})\bar{c}^\epsilon_\gamma, \tag{3.3}$$

which can be written as

$$c^\gamma_\epsilon A^{*\epsilon}_{\alpha k} = c^\gamma_{\alpha/k} + c^\beta_\alpha A^\gamma_{\beta k}. \tag{3.3$'$}$$

If (3.3$'$) is regarded as a system of partial differential equations in the functions c^γ_α, the integrability conditions are easily found to be

$$A^{*\beta}_{\alpha kl} = \bar{c}^\beta_\epsilon c^\gamma_\alpha A^\epsilon_{\gamma kl}, \tag{3.4}$$

where

$$-A^\beta_{\alpha kl} \equiv A^\beta_{\alpha k/l} - A^\beta_{\alpha l/k} + A^\epsilon_{\alpha k}A^\beta_{\epsilon l} - A^\epsilon_{\alpha l}A^\beta_{\epsilon k}, \tag{3.5}$$

with a similar expression for $A^{*\beta}_{\alpha kl}$ in terms of $A^{*\beta}_{\alpha k}$.

The definition (3.5) shows that *for fixed α and β, $A^\beta_{\alpha kl}$ is a tensor skew-symmetric in k, l.* Furthermore, it follows from (3.4) that *for fixed k and l, $A^\beta_{\alpha kl}$ has tensor character with respect to α and β under changes of bases.* This is the reason why we call $A^\beta_{\alpha kl}$ a *double* tensor.

Contraction of $A^\beta_{\alpha kl}$ with respect to α, β gives rise to *the tensor*

$$A_{kl} \equiv A^\alpha_{\alpha kl} = A^\alpha_{\alpha k/l} - A^\alpha_{\alpha l/k}, \tag{3.6}$$

which is invariant under changes of bases.

Thus, with every parallel plane in A_n, there are associated the double tensor $A^\beta_{\alpha kl}$ and the tensor A_{kl}. From our experience with the curvature

tensor of A_n, it is natural to expect the tensors $A^{\beta}_{\alpha kl}$ and A_{kl} to play an important part in the theory of parallel planes in A_n. In §§ 4–6, I shall prove a few theorems concerning these tensors, leaving a fuller investigation to future occasions.

4. Geometric meanings of $A^{\beta}_{\alpha kl}$ and A_{kl}

From (3.1) and (3.5) it follows easily that

$$\lambda^i_{\alpha/kl} - \lambda^i_{\alpha/lk} = -\lambda^i_{\beta} A^{\beta}_{\alpha kl}.$$

On the other hand, we have by Ricci's identity [Eisenhart (1) 12] that

$$\lambda^i_{\alpha/kl} - \lambda^i_{\alpha/lk} = -\lambda^j_{\alpha} B^i_{jkl},$$

where B^i_{jkl} is the curvature tensor of A_n. Hence

$$\lambda^j_{\alpha} B^i_{jkl} = \lambda^i_{\beta} A^{\beta}_{\alpha kl}. \tag{4.1}$$

When the vector λ^i_{α} at any point of A_n is transported parallelly round an infinitesimal circuit in the form of an infinitesimal parallelogram with adjacent edges $d_1 x^i$, $d_2 x^i$, the principal part of the infinitesimal increment in λ^i_{α} is [Eisenhart (1) (10.5)]

$$\Delta\lambda^i_{\alpha} = -\lambda^j_{\alpha} B^i_{jkl} d_1 x^k d_2 x^l,$$

i.e., by (4.1),

$$\Delta\lambda^i_{\alpha} = -\lambda^i_{\beta} A^{\beta}_{\alpha kl} d_1 x^k d_2 x^l. \tag{4.2}$$

Equation (4.2) confirms the fact that $\{\lambda^i_{\alpha}\}$ is a basis of a parallel r-plane. It also gives the following geometric meaning to the double tensor $A^{\beta}_{\alpha kl}$:

$$\text{component of } \Delta\lambda^i_{\alpha} \text{ along } \lambda^i_{\beta} = -A^{\beta}_{\alpha kl} d_1 x^k d_2 x^l. \tag{4.3}$$

A consequence of (4.3) is the following geometric meaning for A_{kl}:

$$\sum_{\alpha} (\text{component of } \Delta\lambda^i_{\alpha} \text{ along } \lambda^i_{\alpha}) = -A_{kl} d_1 x^k d_2 x^l. \tag{4.4}$$

Thus, the left-hand member of (4.4) is independent of the basis; it depends only on the circuit.

5. Bianchi's identity for $A^{\beta}_{\alpha kl}$

Covariant differentiation of (4.1) gives

$$\lambda^j_{\alpha/m} B^i_{jkl} + \lambda^j_{\alpha} B^i_{jkl/m} = \lambda^i_{\beta/m} A^{\beta}_{\alpha kl} + \lambda^i_{\beta} A^{\beta}_{\alpha kl/m},$$

which becomes, on account of (3.1) and then (4.1),

$$A^{\beta}_{\alpha m} \lambda^i_{\gamma} A^{\gamma}_{\beta kl} + \lambda^j_{\alpha} B^i_{jkl/m} = A^{\gamma}_{\beta m} \lambda^i_{\gamma} A^{\beta}_{\alpha kl} + \lambda^i_{\gamma} A^{\gamma}_{\alpha kl/m}.$$

From this equation, two others can be obtained by cyclic permutations of the indices k, l, m. If these three equations are added and the result is simplified by means of Bianchi's identity for B^i_{jkl},

$$B^i_{jkl/m} + B^i_{jlm/k} + B^i_{jmk/l} = 0,$$

we have, since λ^i_γ are independent,

$$A^\gamma_{\alpha kl/m}+A^\gamma_{\alpha lm/k}+A^\gamma_{\alpha mk/l}$$
$$= A^\gamma_{\beta kl}A^\beta_{\alpha m}+A^\gamma_{\beta lm}A^\beta_{\alpha k}+A^\gamma_{\beta mk}A^\beta_{\alpha l}-A^\beta_{\alpha kl}A^\gamma_{\beta m}-A^\beta_{\alpha lm}A^\gamma_{\beta k}-A^\beta_{\alpha mk}A^\gamma_{\beta l}. \tag{5.1}$$

This may be looked upon as Bianchi's identity for $A^\beta_{\alpha kl}$.

6. Some theorems on $A^\beta_{\alpha kl}$ and A_{kl}

THEOREM 6.1. *A parallel plane in A_n is strictly parallel if and only if the double tensor associated with it vanishes identically.*

Proof. We first note that, because of (3.4), the condition $A^\beta_{\alpha kl} = 0$ is invariant under changes of bases. Now, by definition, a parallel plane is *strictly* parallel if it has a basis $\{\lambda^{*i}_\alpha\}$ such that $\lambda^{*i}_{\alpha/k} = 0$ [Walker (8) 139]. For this basis, $A^{*\beta}_{\alpha k} = 0$, and consequently $A^{*\beta}_{\alpha kl} = 0$. Hence $A^\beta_{\alpha kl} = 0$.

Conversely, let us suppose that, for a parallel plane Π, $A^\beta_{\alpha kl} = 0$. Then from the way in which (3.4) was derived from (3.3′), it follows that the integrability conditions of the differential equations

$$(c^\gamma_\epsilon A^{*\epsilon}_{\alpha k} =)\; c^\gamma_{\alpha/k}+c^\beta_\alpha A^\gamma_{\beta k} = 0 \tag{6.1}$$

are identically satisfied. Hence, by choosing as initial values any set of constants $(c^\gamma_\alpha)_0$ with non-zero determinant, we have a solution of (6.1) for c^γ_α such that $A^{*\epsilon}_{\alpha k} = 0$: that is, there exists a basis $\{\lambda^{*i}_\alpha\}$ for which $\lambda^{*i}_{\alpha/k} = 0$. Hence Π is strictly parallel.

THEOREM 6.2. *For a parallel plane Π in A_n, the tensor A_{kl} vanishes identically if and only if Π has a basis for which $A^\alpha_{\alpha k} = 0$.*

Proof. It should be first noted that the vector $A^\alpha_{\alpha k}$ is not invariant under changes of bases.

By (3.6), $A_{kl} = 0$ if $A^\alpha_{\alpha k} = 0$. Conversely, suppose that

$$A_{kl} = A^\alpha_{\alpha k/l}-A^\alpha_{\alpha l/k} = 0.$$

Then a scalar f exists such that $A^\alpha_{\alpha k} = f_{/k}$. From (3.3), we have

$$\begin{aligned} A^{*\alpha}_{\alpha k} &= (c^\gamma_{\alpha/k}+c^\beta_\alpha A^\gamma_{\beta k})\bar{c}^\alpha_\gamma \\ &= c^\gamma_{\alpha/k}\bar{c}^\alpha_\gamma+A^\alpha_{\alpha k} \\ &= [\log\det(c^\beta_\alpha)]_{/k}+f_{/k}. \end{aligned}$$

Thus any change of basis with $\det(c^\beta_\alpha) = \exp(-f)$ will make $A^{*\alpha}_{\alpha k} = 0$, as was to be proved.

COROLLARY 6.2. *For any parallel non-null plane in a Riemannian V_n the tensor A_{kl} vanishes identically.*

Proof. Let $\{\lambda^i_\alpha\}$ be a normal basis of this parallel plane, then $\lambda_{\beta i}\lambda^i_\alpha = \pm 1$ or 0 according as $\beta = \alpha$ or $\beta \neq \alpha$. Transvection of (3.1) by $\lambda_{\alpha i}$ gives

$$0 = \lambda_{\alpha i}\lambda^i_{\alpha/k} = A^\beta_{\alpha k}\lambda^i_\beta\lambda_{\alpha i} = \pm A^\alpha_{\alpha k} \quad (\alpha \text{ not summed}).$$

Therefore $A_{kl} = 0$.

Not every parallel plane in every V_n has $A_{kl} = 0$. Walker [(**8**) 141] has, in fact, given an example of a parallel null 1-plane in a V_3 for which the vector A^1_{1k} in $\lambda^i_{1/k} = A^1_{1k}\lambda^i_1$ is not a gradient. For this parallel 1-plane, $A_{kl} \neq 0$.

7. A_n admitting two parallel planes

The next two theorems generalize some results obtained by Walker (**9**), (**10**) for a V_n.

THEOREM 7.1. *Let A_n admit a parallel $(r+s)$-plane Π and a parallel $(r+t)$-plane Π', which intersect in an r-plane Π^* $(= \Pi \cap \Pi')$. Then there exists in A_n a coordinate system in which $\{\delta^i_\alpha\}$, $\{\delta^i_\alpha, \delta^i_\xi\}$, $\{\delta^i_\alpha, \delta^i_\mu\}$ are bases of Π^*, Π, Π', respectively, where and throughout this section*

$$\alpha, \beta, \gamma \in L = (1,\ldots, r),$$
$$\xi, \eta, \zeta \in M = (r+1,\ldots, r+s),$$
$$\lambda, \mu, \nu \in N = (r+s+1,\ldots, r+s+t),$$
$$\tau, \omega \in T = (r+s+t+1,\ldots, n).$$

This theorem holds for any integer values of r, s, t satisfying $r \geqslant 0$, $s \geqslant 0$ $t \geqslant 0$, $r+s+t \leqslant n$.

Proof. Assume first that none of r, s, t is zero. Let $\{\lambda^i_\alpha, \lambda^i_\xi\}$, $\{\lambda^i_\alpha, \lambda^i_\mu\}$ be bases of Π and Π', respectively. Then the $r+s+t$ vectors λ^i_α, λ^i_ξ, λ^i_μ are independent. Moreover, the r-plane Π^* with basis $\{\lambda^i_\alpha\}$ and the $(r+s+t)$-plane $\Pi \cup \Pi'$ with basis $\{\lambda^i_\alpha, \lambda^i_\xi, \lambda^i_\mu\}$ are also parallel planes. Let us form the following sets of partial differential equations:

$$\text{(i)} \quad \lambda^i_\alpha f_{.i} = 0, \qquad \text{(ii)} \quad \lambda^i_\xi f_{.i} = 0, \qquad \text{(iii)} \quad \lambda^i_\mu f_{.i} = 0,$$

where $f_{.i} = \partial f/\partial x^i$. Then (i), (i, ii), (i, iii), (i, ii, iii) are complete systems associated with the parallel planes Π^*, Π, Π', $\Pi \cup \Pi'$, respectively [cf. remarks on (2.2)].

Let f^ω $(\omega \in T)$ be a complete set of $n-r-s-t$ solutions of (i, ii, iii); (f^ω, f^μ) a complete set of $n-r-s$ solutions of (i, ii); and (f^ω, f^ξ) a complete set of $n-r-t$ solutions of (i, iii). We now prove that *the $n-r$ functions*

f^{ω}, f^{μ}, f^{ξ} *are independent* and therefore form a complete set of $n-r$ solutions of (i).†

Assume that between these functions there exists a functional relation

$$F(f^{\omega}, f^{\mu}, f^{\xi}) = 0. \tag{7.1}$$

Then, since f^{ω}, f^{μ} are independent solutions of (i, ii), we have

$$\frac{\partial F}{\partial f^{\xi}} \neq 0, \tag{7.2}$$

and also $$\lambda^i_\eta f^{\omega}_{.i} = 0, \qquad \lambda^i_\eta f^{\mu}_{.i} = 0. \tag{7.3}$$

It follows from (7.1) and (7.3) that

$$0 = \lambda^i_\eta F_{.i} = \frac{\partial F}{\partial f^{\omega}} \lambda^i_\eta f^{\omega}_{.i} + \frac{\partial F}{\partial f^{\mu}} \lambda^i_\eta f^{\mu}_{.i} + \frac{\partial F}{\partial f^{\xi}} \lambda^i_\eta f^{\xi}_{.i} = \frac{\partial F}{\partial f^{\xi}} \lambda^i_\eta f^{\xi}_{.i}.$$

This, on account of (7.2), requires that $\det(\lambda^i_\eta f^{\xi}_{.i}) = 0$. Therefore, functions ϕ^{η} (not all zero) exist such that

$$\phi^{\eta} \lambda^i_\eta f^{\xi}_{.i} = 0. \tag*{(7.4)$_1$}$$

To these let us add the following consequence of (7.3):

$$\phi^{\eta} \lambda^i_\eta f^{\omega}_{.i} = 0. \tag*{(7.4)$_2$}$$

Now consider the system of $n-r-t$ linear and homogeneous equations in λ^i

$$\lambda^i f^{\xi}_{.i} = 0, \qquad \lambda^i f^{\omega}_{.i} = 0. \tag{7.5}$$

Since f^{ξ}, f^{ω} are $n-r-t$ independent functions, the coefficient matrix $(f^{\xi}_{.i}, f^{\omega}_{.i})$ of equations (7.5) is of rank $n-r-t$. Therefore (7.5) have exactly $r+t$ independent solutions. But, by (i, iii), λ^i_α and λ^i_μ are $r+t$ independent solutions, and by (7.4) $\phi^{\eta}\lambda^i_\eta$ is another solution. Therefore $\phi^{\eta}\lambda^i_\eta$ must be linearly dependent on λ^i_α, λ^i_μ. But this is impossible because λ^i_α, λ^i_ξ, λ^i_μ are independent vectors. Hence the $n-r$ functions $f^{\omega}, f^{\mu}, f^{\xi}$ are independent, as was to be proved.

Now let f^{α} ($\alpha \in L$) be any r functions forming with the preceding $n-r$ functions a set of n independent functions; and let us define a new coordinate system $(\bar{x}^i)$ by $\bar{x}^i = f^i$. Then, since the equations in (i, ii, iii) are invariant under changes of coordinates, we have in $(\bar{x}^i)$

$$\begin{array}{lll} \bar{\lambda}^i_\alpha \bar{x}^{\xi}_{.i} = 0, & \bar{\lambda}^i_\alpha \bar{x}^{\mu}_{.i} = 0, & \bar{\lambda}^i_\alpha \bar{x}^{\omega}_{.i} = 0; \\ & \bar{\lambda}^i_\xi \bar{x}^{\mu}_{.i} = 0, & \bar{\lambda}^i_\xi \bar{x}^{\omega}_{.i} = 0; \\ \bar{\lambda}^i_\mu \bar{x}^{\xi}_{.i} = 0, & & \bar{\lambda}^i_\mu \bar{x}^{\omega}_{.i} = 0, \end{array}$$

† A special case of this fact was used implicitly by Walker in (**10**) 148–9. But it appears to me that the assertion Walker made concerning his (4a), (4b), (4c), and (4d) needs amplification; for, if (4a), (4b), (4c) are assumed, then (4d) must be proved, and this is exactly what I propose to do here.

where $\bar{x}^k_{.i} = \partial\bar{x}^k/\partial\bar{x}^i$. Therefore, the components of $\bar{\lambda}^i_\alpha$, $\bar{\lambda}^i_\xi$, $\bar{\lambda}^i_\mu$ can be partitioned as follows:

$$\bar{\lambda}^i_\alpha = (\bar{\lambda}^\beta_\alpha, 0, 0, 0),$$
$$\bar{\lambda}^i_\xi = (\bar{\lambda}^\beta_\xi, \bar{\lambda}^\eta_\xi, 0, 0),$$
$$\bar{\lambda}^i_\mu = (\bar{\lambda}^\beta_\mu, 0, \bar{\lambda}^\nu_\mu, 0).$$

From these it follows that $\{\delta^i_\alpha\}$, $\{\delta^i_\alpha, \delta^i_\xi\}$, $\{\delta^i_\alpha, \delta^i_\mu\}$ may be taken to replace $\{\bar{\lambda}^i_\alpha\}$, $\{\bar{\lambda}^i_\alpha, \bar{\lambda}^i_\xi\}$, $\{\bar{\lambda}^i_\alpha, \bar{\lambda}^i_\mu\}$ as bases of Π^*, Π, Π', respectively. Thus Theorem 7.1 is proved for the case where none of r, s, t is zero.

Now we can easily verify that the above proof still holds if r is zero, and that it still holds but becomes trivial if one of s, t is zero or both are zero. Hence Theorem 7.1 is completely proved.

THEOREM 7.2. *In an A_n with (symmetric) connexion Γ^i_{jk} the r-plane with basis $\{\delta^i_\alpha\}$ is parallel if and only if $\Gamma^\rho_{\alpha k} = 0$ ($\alpha = 1,..., r$; $\rho = r+1,..., n$).*

Proof. When $\lambda^i_\alpha = \delta^i_\alpha$, the recurrent relations (2.1) become

$$A^\beta_{\alpha k}\delta^i_\beta = \delta^i_{\alpha/k} = \Gamma^i_{\alpha k},$$

which are equivalent to

$$A^\beta_{\alpha k} = \Gamma^\beta_{\alpha k}, \qquad 0 = \Gamma^\rho_{\alpha k}.$$

The first set of equations merely determines the functions $A^\beta_{\alpha k}$; the second set of equations is therefore the condition for the r-plane with basis $\{\delta^i_\alpha\}$ to be parallel [Walker (**9**) 75].

8. Kähler's lemma

We shall need the following lemma† due to Kähler [(**3**) § 2]:

LEMMA. *If two sets of functions f_θ, f_ϕ ($\theta = 1,..., p$; $\phi = p+1,..., p+q$; $p+q \leqslant n$) satisfy $f_{\theta.\phi} = f_{\phi.\theta}$, then there exist functions f, F_θ, F_ϕ such that*

$$F_{\theta.\phi} = 0 = F_{\phi.\theta}, \qquad f_\theta = f_{.\theta}+F_\theta, \qquad f_\phi = f_{.\phi}+F_\phi;$$

and consequently, $$f_{\theta.\phi} = f_{\phi.\theta} = f_{.\theta\phi}.$$

Here, as usual, a dot indicates ordinary partial differentiation.

9. Riemannian V_n admitting a parallel plane

We now consider the case when the A_n is a Riemannian V_n. In V_n, we can talk about the *orthogonality* between vectors, and, in consequence, we have null vectors and partially null planes.

Let Π be a plane in V_n and Π' its *conjugate* plane, i.e. the plane composed of the vectors orthogonal to all the vectors of Π. Then the intersection Π^* of Π and Π' is a null plane, the *null part* of Π (and of Π').

† This lemma was used implicitly by Walker in (**10**) (**17**).

If Π is parallel, so also is Π'. Therefore we have in fact a V_n admitting two parallel planes Π, Π', intersecting in a (parallel) plane Π^*; hence Theorem 7.1 may be applied to this V_n.

As in Theorem 7.1, let Π^*, Π be of dimensions r, $r+s$, respectively. Then Π', being the conjugate of Π, is of dimension $n-r-s$. Therefore, putting $r+t = n-r-s$ in Theorem 7.1, we have

In a V_n admitting a parallel $(r+s)$-plane Π of nullity r (i.e. with null part of dimensions r) there exists a coordinate system in which $\{\delta^i_\alpha\}$, $\{\delta^i_\alpha, \delta^i_\xi\}$, $\{\delta^i_\alpha, \delta^i_\mu\}$ are bases of Π^*, Π, Π', respectively, where Π', Π^* are respectively the conjugate and the null part of Π, and

$$\begin{aligned} \alpha, \beta, \gamma &\in L = (1,..., r), \\ \xi, \eta, \zeta &\in M = (r+1,..., r+s), \\ \lambda, \mu, \nu &\in N = (r+s+1,..., n-r), \\ \tau, \omega &\in L' = (n-r+1,..., n). \end{aligned}$$

I proceed to show how the above coordinate system combined with the orthogonality properties of Π and Π' leads naturally to Walker's canonical form [(9.15) below] for the metric of the most general V_n admitting a parallel $(r+s)$-plane of nullity r. The main difference between my method and Walker's in arriving at the canonical form lies in the initial coordinate system used. The above coordinate system used by me is slightly more general but much easier to establish than the initial coordinate system used by Walker [(**9**) 73–4; (**10**) 148–9].

Let $ds^2 = g_{jk}\,dx^j dx^k$ be the metric of the V_n in the coordinate system described above. Since every vector in Π is orthogonal to every vector in Π', and Π^* is the intersection of Π and Π', each of the vectors in $\{\delta^i_\alpha\}$ is orthogonal to itself and to every other, and the sets of vectors $\{\delta^i_\alpha\}$, $\{\delta^i_\xi\}$, $\{\delta^i_\mu\}$ are mutually orthogonal. Therefore,

$$g_{\alpha\beta} = 0 = g_{\alpha\xi} = g_{\alpha\mu} = g_{\xi\mu}. \tag{9.1}$$

A consequence of these is

$$[\det(g_{\alpha\omega})]^2 \det(g_{\xi\eta}) \det(g_{\mu\nu}) = \det(g_{jk}) \neq 0.$$

Therefore,

$$\det(g_{\alpha\omega}) \neq 0, \qquad \det(g_{\xi\eta}) \neq 0, \qquad \det(g_{\mu\nu}) \neq 0. \tag{9.2}$$

Since Π with basis $\{\delta^i_\alpha, \delta^i_\xi\}$ is parallel, we have, by Theorem 7.2,

$$\Gamma^\phi_{\theta k} = 0 \quad (\theta \in L+M,\ \phi \in N+L').\dagger \tag{9.3}$$

† Similar conditions on Γ^i_{jk} obtained from the fact that each of Π', Π^*, $\Pi \cup \Pi'$ with bases $\{\lambda^i_\alpha, \lambda^i_\mu\}$, $\{\lambda^i_\alpha\}$, $\{\lambda^i_\alpha, \lambda^i_\xi, \lambda^i_\mu\}$ is a parallel plane turn out to be consequences of (9.3) and (9.1); this is as it should be since the existence of the parallel plane Π in V_n implies the existence of the parallel planes Π', Π^*, $\Pi \cup \Pi'$.

It is easy to see that on account of (9.1) and (9.2), conditions (9.3) are equivalent to

$$[\alpha, \theta k] = g_{\alpha i}\,\Gamma^{i}_{\theta k} = g_{\alpha\omega}\,\Gamma^{\omega}_{\theta k} = 0,$$

$$[\mu, \theta k] = g_{\mu i}\,\Gamma^{i}_{\theta k} = g_{\mu\nu}\,\Gamma^{\nu}_{\theta k} + g_{\mu\omega}\,\Gamma^{\omega}_{\theta k} = 0,$$

i.e.
$$g_{\alpha k.\theta} - g_{\theta k.\alpha} = 0, \qquad g_{\mu k.\theta} - g_{\theta k.\mu} = 0,$$

where $[i, jk]$ are the Christoffel symbols of the first kind, and as usual a dot indicates ordinary partial differentiation. By letting

$$k\ (\in L+M+N+L') \quad \text{and} \quad \theta\ (\in L+M)$$

take all the possible values the above equations are seen to reduce to the following:

$$g_{\xi\eta.\alpha} = 0 = g_{\xi\eta.\mu}, \qquad g_{\mu\nu.\alpha} = 0 = g_{\mu\nu.\xi}; \tag{9.4}$$

$$g_{\alpha\omega.\beta} - g_{\beta\omega.\alpha} = 0; \tag{9.5}$$

$$g_{\alpha\omega.\xi} - g_{\xi\omega.\alpha} = 0, \qquad g_{\mu\omega.\alpha} - g_{\alpha\omega.\mu} = 0; \tag{9.6}$$

$$g_{\mu\omega.\xi} - g_{\xi\omega.\mu} = 0. \tag{9.7}$$

Equations (9.4) show that

$$g_{\xi\eta} = g_{\xi\eta}(x^{\zeta}, x^{\tau}), \qquad g_{\mu\nu} = g_{\mu\nu}(x^{\lambda}, x^{\tau}). \tag{9.8}$$

Equations (9.5) show that functions f_{ω} exist such that

$$g_{\alpha\omega} = f_{\omega.\alpha}. \tag{9.9}$$

Substitution of this in (9.6) gives

$$0 = f_{\omega.\alpha\xi} - g_{\xi\omega.\alpha} = (f_{\omega.\xi} - g_{\xi\omega})_{.\alpha},$$

$$0 = g_{\mu\omega.\alpha} - f_{\omega.\alpha\mu} = (g_{\mu\omega} - f_{\omega.\mu})_{.\alpha}.$$

Therefore
$$g_{\xi\omega} = f_{\omega.\xi} + G_{\xi\omega}, \qquad g_{\mu\omega} = f_{\omega.\mu} + G_{\mu\omega}, \tag{9.10}$$

where
$$G_{\xi\omega.\alpha} = 0 = G_{\mu\omega.\alpha}. \tag{9.11}$$

On account of (9.10), equations (9.7) become $G_{\mu\omega.\xi} = G_{\xi\omega.\mu}$. Hence it follows from Kähler's lemma that

$$\left.\begin{aligned} G_{\xi\omega} &= H_{\omega.\xi} + K_{\xi\omega}(x^{\zeta}, x^{\tau}) \\ G_{\mu\omega} &= H_{\omega.\mu} + K_{\mu\omega}(x^{\lambda}, x^{\tau}) \end{aligned}\right\}, \tag{9.12}$$

where, because of (9.11), $\qquad H_{\omega.\alpha} = 0. \qquad (9.13)$

Using (9.1), (9.9), (9.10), (9.12), and (9.13), we now have

$$\begin{aligned} ds^{*2} &= 2g_{\alpha\omega}\,dx^{\alpha}dx^{\omega} + 2g_{\xi\omega}\,dx^{\xi}dx^{\omega} + 2g_{\mu\omega}\,dx^{\mu}dx^{\omega} \\ &= 2f_{\omega.\alpha}\,dx^{\alpha}dx^{\omega} + 2[f_{\omega.\xi} + H_{\omega.\xi} + K_{\xi\omega}(x^{\zeta}, x^{\tau})]\,dx^{\xi}dx^{\omega} + \\ &\qquad + 2[f_{\omega.\mu} + H_{\omega.\mu} + K_{\mu\omega}(x^{\lambda}, x^{\tau})]\,dx^{\mu}dx^{\omega} \\ &= 2(df_{\omega} - f_{\omega.\tau}\,dx^{\tau})\,dx^{\omega} + 2(dH_{\omega} - H_{\omega.\tau}\,dx^{\tau})\,dx^{\omega} + \\ &\qquad + 2K_{\xi\omega}(x^{\zeta}, x^{\tau})\,dx^{\xi}dx^{\omega} + 2K_{\mu\omega}(x^{\lambda}, x^{\tau})\,dx^{\mu}dx^{\omega}. \end{aligned}$$

Therefore

$$\begin{aligned} ds^2 &= ds^{*2}+g_{\xi\eta}\,dx^\xi dx^\eta+g_{\mu\nu}\,dx^\mu dx^\nu+2g_{\tau\omega}\,dx^\tau dx^\omega \\ &= 2d(f_\omega+H_\omega)\,dx^\omega+g_{\xi\eta}(x^\zeta,x^\tau)\,dx^\xi dx^\eta+2K_{\xi\omega}(x^\zeta,x^\tau)\,dx^\xi dx^\omega+ \\ &\qquad +g_{\mu\nu}(x^\lambda,x^\tau)\,dx^\mu dx^\nu+2K_{\mu\omega}(x^\lambda,x^\tau)\,dx^\mu dx^\omega+ \\ &\qquad\qquad +(g_{\omega\tau}-2f_{\omega.\tau}-2H_{\omega.\tau})\,dx^\tau dx^\omega. \end{aligned}$$

Since, by (9.13), (9.9), and (9.2),

$$\det[(f_\omega+H_\omega)_{.\alpha}] = \det(f_{\omega.\alpha}) = \det(g_{\omega\alpha}) \neq 0,$$

the transformation

$$\bar{x}^\alpha = f_{n-r+\alpha}+H_{n-r+\alpha}, \qquad \bar{x}^i = x^i \quad (i \neq \alpha) \tag{9.14}$$

is non-singular. It defines a coordinate system $(\bar{x}^i)$ in which

$$\begin{aligned} ds^2 &= 2\sum_\alpha d\bar{x}^\alpha d\bar{x}^{n-r+\alpha}+\bar{g}_{\xi\eta}(\bar{x}^\zeta,\bar{x}^\tau)\,d\bar{x}^\xi d\bar{x}^\eta+2\bar{g}_{\xi\omega}(\bar{x}^\zeta,\bar{x}^\tau)\,d\bar{x}^\xi d\bar{x}^\omega+ \\ &\quad +\bar{g}_{\mu\nu}(\bar{x}^\lambda,\bar{x}^\tau)\,d\bar{x}^\mu d\bar{x}^\nu+2\bar{g}_{\mu\omega}(\bar{x}^\lambda,\bar{x}^\tau)\,d\bar{x}^\mu d\bar{x}^\omega+\bar{g}_{\tau\omega}(\bar{x}^i)\,d\bar{x}^\tau d\bar{x}^\omega. \end{aligned} \tag{9.15}$$

This is Walker's canonical form [(**10**) 148]. That $\{\delta^i_\alpha\}$, $\{\delta^i_\alpha, \delta^i_\xi\}$, $\{\delta^i_\alpha, \delta^i_\mu\}$ in the coordinate system $(\bar{x}^i)$ are still bases of the parallel planes Π^*, Π, Π' follows immediately from (9.14); this should also have been expected because the coordinate system of Theorem 7.1 has the freedom that the functions f^α in $\bar{x}^i = f^i$ can be chosen arbitrarily but suitably [cf. the proof of Theorem 7.1].

10. Even-dimensional V_n admitting two non-intersecting parallel null $\frac{1}{2}n$-planes

An example of V_4 admitting two non-intersecting parallel null 2-planes was given by Ruse [(**5**) 229], who pointed out that the metric of such V_4 could not be easily obtained from his (and Walker's) canonical form for the metric of the most general V_4 admitting a parallel null 2-plane. In this section, I shall give a characteristic property of even-dimensional V_n admitting two non-intersecting parallel null $\frac{1}{2}n$-planes, and obtain as an application of Theorems 7.1 and 7.2 a canonical form for the metric of such V_n.

We have seen in § 9 how a parallel plane Π in V_n gives rise to the parallel planes Π', $\Pi^* = \Pi \cap \Pi'$, and $\Pi \cup \Pi'$. Let V_n admit two parallel planes Π, Σ, of nullity r, r', and let Π, $\Pi \cup \Pi'$, Σ, $\Sigma \cup \Sigma'$ be of dimensions $r+s$, $n-r$, $r'+s'$, $n-r'$, respectively. Since $\Pi \cup \Pi'$ contains Π^*, and $\Sigma \cup \Sigma'$ contains Σ^*, we have $n-r \geqslant r$, $n-r' \geqslant r'$, i.e. $r \leqslant \frac{1}{2}n$, $r' \leqslant \frac{1}{2}n$. If the planes $\Pi \cup \Pi'$, $\Sigma \cup \Sigma'$ do not intersect, then we must have

$n-r+n-r' \leqslant n$, i.e. $n \leqslant r+r'$. This inequality combined with $r \leqslant \frac{1}{2}n$, $r' \leqslant \frac{1}{2}n$ requires that n be even and $r = \frac{1}{2}n = r'$, $s = 0 = s'$. Hence

THEOREM 10.1. *If a V_n admits two parallel planes Π and Σ, and if $\Pi \cup \Pi'$, $\Sigma \cup \Sigma'$ do not intersect, then n is even and Π, Σ are (non-intersecting) parallel null $\frac{1}{2}n$-planes.*

Let an even-dimensional V_n admit two non-intersecting parallel null $\frac{1}{2}n$-planes Π and Σ. Then, from Theorem 7.1 (with $r = 0$, $s = t = \frac{1}{2}n$), it follows that there exists in V_n a coordinate system in which $\{\delta^i_\xi\}$, $\{\delta^i_\mu\}$ are bases of Π, Σ, respectively, where and throughout this section

$$\xi, \eta \in M = (1,..., \tfrac{1}{2}n); \qquad \mu, \nu \in N = (\tfrac{1}{2}n+1,..., n).$$

In this coordinate system we have, by Theorem 7.2,

$$\Gamma^{\mu}_{\xi k} = 0 = \Gamma^{\xi}_{\mu k}. \tag{10.1}$$

Moreover, since Π and Σ are null,

$$g_{\xi\eta} = 0 = g_{\mu\nu}. \tag{10.2}$$

On account of these, equations (10.1) are equivalent to

$$[\eta, \xi k] = g_{\eta i}\,\Gamma^{i}_{\xi k} = g_{\eta\mu}\,\Gamma^{\mu}_{\xi k} = 0,$$

$$[\nu, \mu k] = g_{\nu i}\,\Gamma^{i}_{\mu k} = g_{\nu\xi}\,\Gamma^{\xi}_{\mu k} = 0,$$

i.e. $$g_{\eta\mu.\xi} - g_{\xi\mu.\eta} = 0, \qquad g_{\nu\xi.\mu} - g_{\mu\xi.\nu} = 0.$$

Therefore functions f_μ, f_ξ exist such that

$$f_{\mu.\xi} = g_{\mu\xi} = g_{\xi\mu} = f_{\xi.\mu}.$$

Hence, by Kähler's lemma,

$$g_{\xi\mu} = f_{\xi.\mu} = f_{.\xi\mu},$$

where f is some function of x^i. From these and (10.2) it follows that the metric of this V_n is

$$ds^2 = 2f_{.\xi\mu}\,dx^\xi dx^\mu.$$

Hence we have the following theorem:

THEOREM 10.2. *In an even-dimensional Riemannian V_n admitting two non-intersecting parallel null $\frac{1}{2}n$-planes, there exists a coordinate system in which the metric of V_n has the form*

$$ds^2 = 2f_{.\xi\mu}\,dx^\xi dx^\mu \quad (\xi = 1,..., \tfrac{1}{2}n;\ \mu = \tfrac{1}{2}n+1,..., n), \tag{10.3}$$

and $\{\delta^i_\xi\}$, $\{\delta^i_\mu\}$ are bases of these parallel null $\frac{1}{2}n$-planes.

It is interesting to observe that (10.3) is the well-known Kähler metric in real variables (**3**).†

† Theorem 10.2 has been given in a note by E. M. Patterson [*J. of London Math. Soc.* 28 (1953) 260–9] published since this paper was written.

REFERENCES

1. E. P. Eisenhart, 'Non-Riemannian geometry', *American Math. Soc. Coll. Pub.*, vol. 8 (New York, 1927).
2. —— 'Fields of parallel vectors in Riemannian space', *Ann. of Math.* 39 (1938), 316–21.
3. E. Kähler, 'Über eine bemerkenswerte Hermitische Metrik', *Abh. Math. Sem. Hamburg*, 9 (1933), 173–86.
4. E. M. Patterson, 'On symmetric recurrent tensors of the second order', *Quart. J. of Math.* (Oxford) (2) 2 (1951), 151–8.
5. H. S. Ruse, 'On parallel fields of planes in a Riemannian space', ibid. (1) 20 (1949), 218–34.
6. —— 'Parallel planes in a Riemannian V_n', *Proc. Royal Soc. Edinburgh*, Section A, 58 (1950), 78–92.
7. J. A. Schouten and W. v. d. Kulk, *Pfaff's Problem and its Generalizations* (Oxford, 1949).
8. A. G. Walker, 'On parallel fields of partially null vector spaces', *Quart. J. of Math.* (Oxford) (1) 20 (1949), 135–45.
9. —— 'Canonical form for a Riemannian space with a parallel field of null planes', ibid. (2) 1 (1950), 67–79.
10. —— 'Canonical forms (II): Parallel partially null planes', ibid. 147–52.

CORRIGENDUM (received 24 September 1953)

THE CHARACTERIZATION OF GENERALIZED CONVEX FUNCTIONS

By F. F. BONSALL (*Newcastle*)

[*Quart. J. of Math.* (Oxford) (2) 1 (1950), 100–11]

The assertion of Theorem 3 Corollary (ii), repeated in Theorem 4, that f has a second derivative p.p. is false. In fact Corollary (ii) does not follow from Corollary (i), since the countable set where the first derivative does not exist may have a derived set with positive measure. This error does not affect any other result in the paper.

EIGENFUNCTION EXPANSIONS ASSOCIATED WITH PARTIAL DIFFERENTIAL EQUATIONS (IV)

By E. C. TITCHMARSH (*Oxford*)

[Received 14 February 1953]

1. In the previous papers of this series† I have considered the problem of the eigenfunction expansion associated with the partial differential equation

$$\nabla^2\phi+\{\lambda-q(x,y)\}\phi = 0, \tag{1.1}$$

the region in which it holds being the whole (x,y)-plane. The object of this paper is to extend the theory to more general partial differential equations.

Consider the equation

$$\frac{\partial}{\partial x}\left\{p(x,y)\frac{\partial\phi}{\partial x}\right\}+\frac{\partial}{\partial y}\left\{p(x,y)\frac{\partial\phi}{\partial y}\right\}+\{\lambda s(x,y)-q(x,y)\}\phi = 0, \tag{1.2}$$

where p, q, and s are given functions of x and y, and λ is the eigenvalue parameter. This is self-adjoint in the sense that, if the left-hand side is denoted by $L\phi$, then

$$\iint \phi_1\, L\phi_2\, dxdy = \iint \phi_2\, L\phi_1\, dxdy,$$

for suitable functions ϕ_1 and ϕ_2. This is easily verified by integration by parts.

Now suppose that $p(x,y)$ is positive and has continuous partial derivatives up to the second order. Let

$$\psi(x,y) = \{p(x,y)\}^{\frac{1}{2}}\phi(x,y).$$

Then (1.2) reduces to

$$\nabla^2\psi+\{\lambda\, s_1(x,y)-q_1(x,y)\}\psi = 0, \tag{1.3}$$

where
$$q_1(x,y) = \frac{1}{2p}\left(\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}\right)-\frac{1}{4p^2}\left\{\left(\frac{\partial p}{\partial x}\right)^2+\left(\frac{\partial p}{\partial y}\right)^2\right\}+\frac{q}{p},$$

$$s_1(x,y) = s(x,y)/p(x,y).$$

This is of the same form as before, but with $p(x,y) = 1$. Thus we shall actually consider

$$\nabla^2\psi+\{\lambda\, s(x,y)-q(x,y)\}\psi = 0. \tag{1.4}$$

We shall suppose that q and s are continuous functions of x and y.

The method by which (1.1) was treated in my paper (**5**) consisted in

† See list of references at the end.

obtaining the theory in the case where the (x, y)-region is a square, and then considering the limiting case as the side of the square tends to infinity. I shall indicate briefly how to carry out the corresponding programme in the case of (1.4).

2. We first want to obtain the theory of (1.4) for a finite region of the (x, y)-plane, which we may take to be the square $(0, \pi; 0, \pi)$, with $\psi = 0$ on the boundary. Suppose that in the square

$$s(x,y) \geqslant s_0 > 0, \qquad |q(x,y)| \leqslant q_1.$$

We derive the theory of (1.4) from that of

$$\nabla^2\psi + \{\kappa - q(x,y)\}\psi = 0 \tag{2.1}$$

with the same region and boundary condition, and it is convenient that $\kappa = 0$ should not be an eigenvalue in this problem. To ensure this, write (1.4) as

$$\nabla^2\psi + \{\lambda' s(x,y) - Q(x,y)\}\psi = 0,$$

where $\qquad \lambda' = \lambda + C, \qquad Q(x,y) = C\,s(x,y) + q(x,y).$

Then $q(x, y)$ is replaced by $Q(x, y)$ in (2.1), and

$$Q(x,y) \geqslant Cs_0 - q_1.$$

If the eigenvalues and normalized eigenfunctions of (2.1) (with Q) are κ_n, χ_n, then

$$\kappa_n = \int\limits_0^\pi \int\limits_0^\pi \left\{\left(\frac{\partial \chi_n}{\partial x}\right)^2 + \left(\frac{\partial \chi_n}{\partial y}\right)^2 + Q(x,y)\chi_n^2\right\} dxdy$$

$$\geqslant (Cs_0 - q_1) \int\limits_0^\pi \int\limits_0^\pi \chi_n^2 \, dxdy = Cs_0 - q_1 > 0$$

if $C > q_1/s_0$. We can therefore suppose without loss of generality that this is true for the original equation (2.1).

Consider, as in (**5**), the equation

$$\nabla^2\Phi + \{\lambda s(x,y) - q(x,y)\}\Phi = f(x,y), \tag{2.2}$$

where f is any function of L^2. We want to construct a solution of this equation which vanishes round the boundary of the square. For this purpose we define a sequence of functions $\Phi_n(x, y)$, which vanish round the boundary, and satisfy

$$\left.\begin{aligned} \{\nabla^2 - q(x,y)\}\Phi_1 &= f(x,y) \\ \{\nabla^2 - q(x,y)\}\Phi_n &= f(x,y) - \lambda s(x,y)\Phi_{n-1} \quad (n = 2, 3,\ldots) \end{aligned}\right\}. \tag{2.3}$$

Let $g(x,y,\xi,\eta,\kappa)$ be the Green's function associated with (2.1). The solutions of (2.3) are given by

$$\left.\begin{aligned} \Phi_1(x,y) &= -\int_0^\pi\int_0^\pi g(x,y,\xi,\eta,0)f(\xi,\eta)\,d\xi d\eta \\ \Phi_n(x,y) &= -\int_0^\pi\int_0^\pi g(x,y,\xi,\eta,0)\{f(\xi,\eta)-\lambda s(\xi,\eta)\Phi_{n-1}(\xi,\eta)\}\,d\xi d\eta \end{aligned}\right\}. \tag{2.4}$$

To prove that the sequence converges, suppose that

$$|\Phi_n(x,y)-\Phi_{n-1}(x,y)| \leqslant \delta_n$$

throughout the square. Then

$$\begin{aligned} |\Phi_{n+1}(x,y)-\Phi_n(x,y)| &\leqslant |\lambda|\delta_n \int_0^\pi\int_0^\pi |g(x,y,\xi,\eta,0)|s(\xi,\eta)\,d\xi d\eta \\ &\leqslant \tfrac{1}{2}\delta_n \end{aligned}$$

if $|\lambda|$ is small enough, since the integral on the right-hand side is bounded. It follows as in (**5**) that $\Phi_n(x,y)$ tends to a limit $\Phi(x,y)$, which is a solution of (2.2), if $|\lambda|$ is small enough.

We also have

$$\Phi_n(x,y) = -\int_0^\pi\int_0^\pi G_n(x,y,\xi,\eta,\lambda)f(\xi,\eta)\,d\xi d\eta,$$

where

$$\begin{aligned} G_2(x,y,\xi,\eta,\lambda) &= g(x,y,\xi,\eta,0)+ \\ &\quad +\lambda\int_0^\pi\int_0^\pi g(x,y,u,v,0)g(u,v,\xi,\eta,0)s(u,v)\,dudv, \end{aligned}$$

and generally

$$\begin{aligned} &G_n(x,y,\xi,\eta,\lambda) \\ &\quad = g(x,y,\xi,\eta,0)+\lambda\int_0^\pi\int_0^\pi g(x,y,u,v,0)G_{n-1}(u,v,\xi,\eta,\lambda)s(u,v)\,dudv. \end{aligned}$$

It can be shown as before that $G_n(x,y,\xi,\eta,\lambda)$ converges to a limit $G(x,y,\xi,\eta,\lambda)$ if $|\lambda|$ is small enough. This defines the Green's function for the equation (1.4) in the square. We also have

$$\Phi(x,y,\lambda) = -\int_0^\pi\int_0^\pi G(x,y,\xi,\eta,\lambda)f(\xi,\eta)\,d\xi d\eta$$

if $|\lambda|$ is small enough.

We next prove as in (5) that

$$(\lambda-\lambda')\int_0^\pi\int_0^\pi G(x,y,u,v,\lambda)G(\xi,\eta,u,v,\lambda')s(u,v)\,dudv$$
$$= G(x,y,\xi,\eta,\lambda)-G(x,y,\xi,\eta,\lambda'),$$

and that, if $G^{(n)} = \partial^n G/\partial\lambda^n$,

$$G^{(n)}(x,y,\xi,\eta,\lambda) = n\int_0^\pi\int_0^\pi G^{(n-1)}(x,y,u,v,\lambda)G(\xi,\eta,u,v,\lambda)s(u,v)\,dudv.$$

From this we deduce that, if $\lambda = \mu+i\nu$,

$$\int_0^\pi\int_0^\pi |\Phi(x,y,\lambda)|^2 s(x,y)\,dxdy \leqslant \frac{1}{\nu^2}\int_0^\pi\int_0^\pi \frac{\{f(x,y)\}^2}{s(x,y)}\,dxdy.$$

Proceeding as before, we obtain the inequality

$$|G^{(n)}(x,y,\xi,\eta,\lambda)|^2$$
$$\leqslant \frac{(n!)^2}{\nu^{2n-2}}\int_0^\pi\int_0^\pi |G(x,y,u,v,\lambda)|^2 s(u,v)\,dudv\int_0^\pi\int_0^\pi |G(\xi,\eta,u,v,\lambda)|^2 s(u,v)\,dudv.$$

Suppose now that $G(x,y,\xi,\eta,\lambda)$ is regular in the circle $|\lambda| < R$. Let $\lambda' = \mu'+i\nu'$ be a point in the upper half of the circle, and expand $G(x,y,\xi,\eta,\lambda)$ in powers of $\lambda-\lambda'$. As in (5), the radius of convergence of the expansion is at least ν', i.e. the expansion converges at least in the circle with centre λ' touching the real axis. It is easily seen in this way that $G(x,y,\xi,\eta,\lambda)$ can be continued analytically over the whole λ-plane, except for the real axis from R to ∞.

The theory now proceeds in the same manner as before. The function $G(x,y,\xi,\eta,\lambda)$ has poles at certain points λ_n, the residues being multiples of the eigenfunctions $\psi_n(x,y)$. These are orthogonal in the sense that

$$\int_0^\pi\int_0^\pi \psi_m(x,y)\psi_n(x,y)s(x,y)\,dxdy = 0 \quad (m \neq n),$$

and they may be normalized so that this integral is 1 if $m = n$. Corresponding to any function $f(x,y)$, we write

$$c_n = \int_0^\pi\int_0^\pi f(x,y)\psi_n(x,y)s(x,y)\,dxdy.$$

The expansion formula is then

$$f(x,y) = \sum_{n=0}^{\infty} c_n\psi_n(x,y),$$

and the Parseval formula is

$$\int_0^\pi \int_0^\pi \{f(x,y)\}^2 s(x,y)\,dxdy = \sum_{n=0}^\infty c_n^2.$$

3. To define the Green's function in the case of the whole plane, we proceed as in § 11 of (**5**). Extending the above analysis by a change of variable to the square $(-b,b;-b,b)$, we distinguish the functions relating to this square by a suffix b. We have

$$\left| \int_{-b}^{b} \int_{-b}^{b} G_b(x_0,y_0,x,y,\lambda) f(x,y)\,dxdy \right| \leqslant \frac{K}{|\nu|} \left\{ \int_{-b}^{b} \int_{-b}^{b} \frac{|f(x,y)|^2}{s(x,y)}\,dxdy \right\}^{\frac{1}{2}},$$

where K is independent of b. Taking $f(x,y) = \bar{G}_b(x_0,y_0,x,y,\lambda)s(x,y)$, we obtain

$$\int_{-b}^{b} \int_{-b}^{b} |G_b(x_0,y_0,x,y,\lambda)|^2 s(x,y)\,dxdy \leqslant \frac{K^2}{\nu^2}.$$

From this point the proof proceeds as before. We obtain a Green's function $G(x_0,y_0,x,y,\lambda)$ for the whole plane with the same properties as before, except that (11.10) of (**5**) is replaced by

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(x,y,\xi,\eta,\lambda)|^2 s(x,y)\,dxdy < \infty.$$

4. The problem of the uniqueness of the Green's function of (1.1) has been considered by me (**3**) and by Sears (**2**), the final result being that it is unique if $q(x,y) \geqslant -Q(r)$, where $r = \sqrt{(x^2+y^2)}$, $Q(r)$ is positive,

$$\int^{\infty} \{Q(r)\}^{-\frac{1}{2}}\,dr$$

is divergent, and some subsidiary conditions are satisfied. It will now be proved that the Green's function of (1.4) is unique if there are functions $S(r)$, $Q(r)$ such that

$$s(x,y) \geqslant S(r), \qquad q(x,y) \geqslant -Q(r),$$

where $S(r) \geqslant 0$, $Q(r) \geqslant \delta > 0$, $S'(r)$ and $Q'(r)$ are continuous,

$$S'(r) = O\{S(r)\}, \qquad Q'(r) = O[\{Q(r)\}^{\frac{3}{2}}],$$

$Q(r)/S(r) \to \infty$ as $r \to \infty$, and

$$\int^{\infty} S(r)\{Q(r)\}^{-\frac{1}{2}}\,dr$$

is divergent.

The proof is a straightforward extension of that which applies to

(1.1). Suppose that $G_1(x,y,\xi,\eta,\lambda)$ and $G_2(x,y,\xi,\eta,\lambda)$ both satisfy the conditions for a Green's function of (1.4), and let

$$g(x,y) = G_1(x,y,\xi,\eta,\lambda) - G_2(x,y,\xi,\eta,\lambda)$$

for given ξ, η, and λ with $\lambda = \mu + i\nu$, $\nu > 0$. If (r,θ) are the polar coordinates corresponding to (x,y), we write $g(x,y) = g[r,\theta]$, and similarly for other functions. Then

$$\int_0^\infty \int_0^{2\pi} |g[r,\theta]|^2 s[r,\theta] r \, dr d\theta < \infty. \tag{4.1}$$

As in § 4 of (**3**),

$$\iint_{r \leqslant R} |g[r,\theta]|^2 s[r,\theta] r \, dr d\theta = \frac{1}{\nu} \int_0^{2\pi} \operatorname{im}\{g[R,\theta]\bar{g}_R[R,\theta]\} R \, d\theta, \tag{4.2}$$

where the suffix denotes partial differentiation. Let $f(r)$ be any integrable function, and

$$F(R) = \int_0^R f(r) \, dr.$$

Multiplying (4.2) by $f(r)$ and integrating over $r \leqslant T$, we obtain, as $T \to \infty$,

$$\iint_{r \leqslant T} \{F(T) - F(r)\} |g[r,\theta]|^2 s[r,\theta] r \, dr d\theta = \frac{1}{\nu} \iint_{r \leqslant T} \operatorname{im}(g\bar{g}_r) f(r) r \, dr d\theta$$

$$= O\left\{ \iint_{r \leqslant T} |g|^2 s r \, dr d\theta \iint_{r \leqslant T} |g_r|^2 f^2 s^{-1} r \, dr d\theta \right\}^{\frac{1}{2}}$$

$$= O\left\{ \iint_{r \leqslant T} g_r^2[r,\theta] f^2(r) \{S(r)\}^{-1} r \, dr d\theta \right\}^{\frac{1}{2}}, \tag{4.3}$$

by (4.1).

If $g[r,\theta]$ is not identically zero, there are positive constants K, R_0 such that

$$\iint_{r \leqslant R} |g[r,\theta]|^2 s[r,\theta] r \, dr d\theta \geqslant K \quad (R \geqslant R_0).$$

Multiplying by $f(R)$ and integrating over (R_0, T), we obtain

$$K\{F(T) - F(R_0)\} \leqslant \iint_{r \leqslant T} \{F(T) - F(r)\} |g[r,\theta]|^2 s[r,\theta] r \, dr d\theta.$$

If $f(r)$ can be chosen so that the integral (4.3) is bounded, while $F(T) \to \infty$, we shall obtain a contradiction.

If $\phi(r)$ is any positive function with a continuous derivative,

$$\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)g\nabla^2\bar{g}\,dxdy$$

$$= -\iint\limits_{r\leqslant T}\left[\frac{\partial\bar{g}}{\partial x}\frac{\partial}{\partial x}\left\{\left(1-\frac{r}{T}\right)\phi g\right\}+\frac{\partial\bar{g}}{\partial y}\frac{\partial}{\partial y}\left\{\left(1-\frac{r}{T}\right)\phi g\right\}\right]dxdy$$

$$= -\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)(|g_x|^2+|g_y|^2)\,dxdy+\frac{1}{T}\iint\limits_{r\leqslant T}\phi(r)g\bar{g}_r\,dxdy -$$

$$-\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi'(r)g\bar{g}_r\,dxdy.$$

Since $\nabla^2\bar{g}=(q-\bar{\lambda}s)\bar{g}$, we obtain, on taking real parts,

$$\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)(|g_x|^2+|g_y|^2)\,dxdy$$

$$= -\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)(q-\mu s)|g|^2r\,drd\theta+$$

$$+\frac{1}{2T}\iint\limits_{r\leqslant T}\phi(r)(g\bar{g}_r+\bar{g}g_r)r\,drd\theta-\tfrac{1}{2}\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi'(r)(g\bar{g}_r+\bar{g}g_r)r\,drd\theta.$$

Since $$|g_x|^2+|g_y|^2=|g_r|^2+r^{-2}|g_\theta|^2,$$
it follows that

$$\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)|g_r|^2r\,drd\theta\leqslant\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)(\mu s-q)|g|^2r\,drd\theta+$$

$$+\frac{1}{2T}\iint\limits_{r\leqslant T}\phi(r)(g\bar{g}_r+\bar{g}g_r)r\,drd\theta-\tfrac{1}{2}\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi'(r)(g\bar{g}_r+\bar{g}g_r)r\,drd\theta$$

$$=J_1+J_2+J_3,$$

say. Now

$$J_1\leqslant\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\phi(r)\left\{|\mu|+\frac{Q(r)}{S(r)}\right\}|g|^2sr\,drd\theta=O(1)$$

if $\phi(r)=S(r)/Q(r)$. Also

$$J_2=\frac{S(T)}{2Q(T)}\int\limits_0^{2\pi}|g[T,\theta]|^2\,d\theta-\frac{1}{2T}\iint\limits_{r\leqslant T}|g|^2\left\{\frac{S(r)}{Q(r)}+\frac{rS'(r)}{Q(r)}-\frac{rS(r)Q'(r)}{Q^2(r)}\right\}drd\theta$$

by integration by parts. The last integral is

$$O\left\{\iint\limits_{r\leqslant T}|g|^2S(r)r\,drd\theta\right\}=O\left\{\iint\limits_{r\leqslant T}|g|^2s[r,\theta]r\,drd\theta\right\}=O(1).$$

The first term on the right is bounded for some arbitrarily large values of T, since on omitting $Q(T)$, replacing $S(T)$ by $s[T,\theta]$, multiplying by T and integrating, we obtain a bounded integral, by (4.1). Hence J_2 is bounded for such values of T.

In J_3 we use the inequality $2|ab| \leqslant a^2+b^2$, with

$$a^2 = |g_r|^2SQ^{-1}, \qquad b^2 = |g|^2(S'Q+Q'S)^2S^{-1}Q^{-3}.$$

We obtain

$$J_3 \leqslant \frac{1}{2}\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\left\{\frac{|g_r|^2S}{Q}+\frac{|g|^2(S'Q+Q'S)^2}{SQ^3}\right\}r\,drd\theta.$$

The last term is $$O\left(\iint\limits_{r\leqslant T}|g|^2Sr\,drd\theta\right) = O(1)$$

as before. Hence

$$\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\frac{S}{Q}|g_r|^2r\,drd\theta \leqslant \frac{1}{2}\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\frac{S}{Q}|g_r|^2r\,drd\theta+O(1),$$

$$\iint\limits_{r\leqslant T}\left(1-\frac{r}{T}\right)\frac{S}{Q}|g_r|^2r\,drd\theta = O(1).$$

On taking $f(r) = S(r)\{Q(r)\}^{-\frac{1}{2}}$, the result stated follows.

5. The next problem to be considered is that of the discreteness of the spectrum of (1.4). A condition that the spectrum of (1.1) should be discrete is that $q(x,y) \to \infty$ as $r = \sqrt{(x^2+y^2)} \to \infty$ [see (**4**)]. The corresponding condition for (1.4) is that $q(x,y)/s(x,y) \to \infty$. We shall prove this, subject to the condition that $s(x,y) = O(r^k)$ for some positive k. It seems likely that the result holds more generally, but the method used here (comparison with Bessel-function cases) requires this condition. A possible application of quite a different method has been noted by Friedrichs (**1**).

Consider first a finite region, which may be taken to be the circle $r \leqslant a$, with boundary condition $\psi = 0$ on $r = a$. In this case the spectrum of (1.4) is discrete. Let the eigenvalues be $\lambda_{n,a}$. We then have the following theorems:

(i) *Let $a < b$, and let $\lambda_{n,b}$ denote the eigenvalues in the corresponding problem with radius b. Then $\lambda_{n,a} \geqslant \lambda_{n,b}$ for each n.*

(ii) *Let $\mu_{n,a}$ denote the eigenvalues in the problem with the same a and $s(x,y)$, but $q(x,y)$ replaced by $Q(x,y)$, where $Q(x,y) \leqslant q(x,y)$. Then $\lambda_{n,a} \geqslant \mu_{n,a}$ for each n.*

(iii) *Let $\mu_{n,a}$ denote the eigenvalues in the problem with the same a and $q(x,y)$, but $s(x,y)$ replaced by $S(x,y)$, where $S(x,y) \geqslant s(x,y)$. Then $\lambda_{n,a} \geqslant \mu_{n,a}$ if $\mu_{n,a} \geqslant 0$.*

Theorem (i) is a case of Theorem 3 of Courant and Hilbert, *Methoden der Math. Physik* (I), ch. VI § 2, and Theorems (ii) and (iii) are cases of Theorem 7 of the same section.

We also require the following theorem.

(iv) *Let λ_n denote the eigenvalues in the problem in which*

$$\left.\begin{aligned} \nabla^2\psi+\lambda s(x,y)\psi &= 0 \quad (0<r<a) \\ \nabla^2\psi+(\lambda-\gamma)s(x,y)\psi &= 0 \quad (a \leqslant r \leqslant b) \end{aligned}\right\}, \tag{5.1}$$

where γ is a positive constant, and $\psi = 0$ on $r = b$. Let μ_n denote the eigenvalues in the corresponding problem in which $s(x,y)$ is replaced by $S(x,y)$, where $S(x,y) \geqslant s(x,y)$. Then $\lambda_n \geqslant \mu_n$ provided that $\mu_n \geqslant \gamma$.

If $\gamma = 0$, this is a case of (iii). To prove it generally, we recall the following formulae in the theory of (1.4). Let $\psi_n(x,y)$ be the eigenfunction corresponding to λ_n in the problem of (5.1), let $f(x,y)$ be any function of L^2, and

$$c_n = \iint\limits_{r \leqslant b} f\psi_n s\, dxdy.$$

If f has partial derivatives up to the second order, let

$$f^*(x,y) = (qf-\nabla^2 f)/s, \qquad c_n^* = \iint\limits_{r \leqslant b} f^*\psi_n s\, dxdy.$$

If $f = 0$ on $r = b$, we have

$$\iint\limits_{r \leqslant b} (f\nabla^2\psi_n - \psi_n \nabla^2 f)\, dxdy = 0,$$

i.e.
$$\iint\limits_{r \leqslant b} \{f(q-\lambda_n s)\psi_n - \psi_n(qf-sf^*)\}\, dxdy = 0,$$

whence
$$c_n^* = \lambda_n c_n.$$

Let
$$D(f) = \iint\limits_{r \leqslant b} (f_x^2+f_y^2+qf^2)\, dxdy.$$

Then
$$D(f) = \iint\limits_{r \leqslant b} (qf^2 - f\nabla^2 f)\, dxdy = \iint\limits_{r \leqslant b} ff^*s\, dxdy.$$

Hence the Parseval formula gives

$$D(f) = \sum_{n=0}^{\infty} c_n c_n^* = \sum_{n=0}^{\infty} \lambda_n c_n^2.$$

Now apply these formulae to (5.1), and let μ_n, $\chi_n(x,y)$, d_n, $\mathbf{D}(f)$ refer similarly to the problem involving $S(x,y)$. Then

$$D(f) = \iint\limits_{r\leqslant b} (f_x^2+f_y^2)\,dxdy + \gamma \iint\limits_{a\leqslant r\leqslant b} sf^2\,dxdy,$$

$$\mathbf{D}(f) = \iint\limits_{r\leqslant b} (f_x^2+f_y^2)\,dxdy + \gamma \iint\limits_{a\leqslant r\leqslant b} Sf^2\,dxdy.$$

Hence
$$\mathbf{D}(f) - D(f) = \gamma \iint\limits_{a\leqslant r\leqslant b} (S-s)f^2\,dxdy.$$

First let $f = \psi_0$. Then $D(f) = \lambda_0$, and

$$\mathbf{D}(f) = \sum_{n=0}^{\infty} \mu_n d_n^2 \geqslant \mu_0 \sum_{n=0}^{\infty} d_n^2 = \mu_0 \iint\limits_{r\leqslant b} S\psi_0^2\,dxdy$$

$$= \mu_0 + \mu_0 \iint\limits_{r\leqslant b} (S-s)\psi_0^2\,dxdy.$$

Hence

$$\lambda_0 - \mu_0 \geqslant \mu_0 \iint\limits_{r\leqslant b} (S-s)\psi_0^2\,dxdy - \gamma \iint\limits_{a\leqslant r\leqslant b} (S-s)\psi_0^2\,dxdy.$$

This is not negative if $\mu_0 \geqslant \gamma$, and so the result follows in the case $n = 0$.

Next let $f = c_0\psi_0 + c_1\psi_1$, where $c_0^2+c_1^2 = 1$, and where the ratio of c_0 to c_1 is determined so that $d_0 = 0$. Then

$$D(f) = \lambda_0 c_0^2 + \lambda_1 c_1^2 \leqslant \lambda_1(c_0^2+c_1^2) = \lambda_1,$$

$$\mathbf{D}(f) = \sum_{n=1}^{\infty} \mu_n d_n^2 \geqslant \mu_1 \sum_{n=1}^{\infty} d_n^2 = \mu_1 \iint\limits_{r\leqslant b} Sf^2\,dxdy,$$

and it follows as before that $\lambda_1 \geqslant \mu_1$ if $\mu_1 \geqslant 0$; and so generally.

6. Now consider (1.4), and suppose that $q(x,y)/s(x,y) \to \infty$. This implies that $q(x,y) \geqslant 0$ except possibly in a bounded region, and, as in § 2, by a change of the λ-origin we can suppose that $q(x,y) \geqslant 0$ for all (x,y). Let γ be a given positive number, and let $q(x,y) \geqslant \gamma s(x,y)$ for $r \geqslant a$.

Let $b > a$, and let $\lambda_{n,b}$ denote the eigenvalues in the problem of (1.4) for $r \leqslant b$, with $\psi = 0$ on $r = b$. Let $\mu_{n,b}$ denote the eigenvalues in the problem of (5.1) with $\psi = 0$ on $r = b$. Then, by (ii) of § 5, $\lambda_{n,b} \geqslant \mu_{n,b}$ for each n.

Let $s(x,y) \leqslant S(r)$. Let $\nu_{n,b}$ denote the eigenvalues in the problem of

$$\left.\begin{aligned} \nabla^2\psi+\lambda S(r)\psi &= 0 \quad (r < a) \\ \nabla^2\psi+(\lambda-\gamma)S(r)\psi &= 0 \quad (a \leqslant r \leqslant b) \end{aligned}\right\}. \tag{6.1}$$

Then, by (iv) of § 5, $\mu_{n,b} \geqslant \nu_{n,b}$ provided that $\nu_{n,b} \geqslant \gamma$.

Suppose that the number N of numbers $\nu_{n,b}$ not exceeding γ is independent of b, if b is large enough. Then, for large enough b,

$$\nu_{0,b} \leqslant \nu_{1,b} \leqslant \ldots \leqslant \nu_{N,b} \leqslant \gamma < \nu_{N+1,b}, \ldots.$$

Hence $\mu_{N+1,b}, \mu_{N+2,b},\ldots$ are greater than γ, and so the number of eigenvalues $\mu_{n,b}$ which do not exceed γ is at most N. This is therefore also true of the eigenvalues $\lambda_{n,b}$.

As b tends steadily to infinity, each $\lambda_{n,b}$ decreases, by (i) of § 5, and it follows as in § 2 of (4) that the spectrum in the problem of (1.4) in the whole plane is discrete for $\lambda < \gamma$, there being at most N eigenvalues less than γ. Since γ is arbitrary, the spectrum in this problem is actually discrete over the whole λ-range.

The problem is thus reduced to that of the eigenvalues of (6.1). In polar coordinates, these equations are

$$\frac{\partial^2\psi}{\partial r^2}+\frac{1}{r}\frac{\partial\psi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2}+\lambda S(r)\psi = 0 \quad (r < a), \text{ etc.} \tag{6.2}$$

The eigenfunctions are of the form

$$\psi = \chi(r)\,{\cos \atop \sin}\,n\theta \quad (n = 0, 1,\ldots),$$

where
$$\frac{d^2\chi}{dr^2}+\frac{1}{r}\frac{d\chi}{dr}-\frac{n^2}{r^2}\chi+\lambda S\chi = 0 \quad (r < a), \text{ etc.}$$

On putting
$$\chi = S^{-\frac{1}{4}}r^{-\frac{1}{2}}\omega, \qquad \rho = \int\limits_0^r \{S(t)\}^{\frac{1}{2}}\,dt$$

these equations become

$$\left.\begin{aligned} \frac{d^2\omega}{d\rho^2}+\left(\lambda-\frac{S''}{4S^2}+\frac{5}{16}\frac{S'^2}{S^3}-\frac{n^2-\frac{1}{4}}{Sr^2}\right)\omega &= 0 \quad (\rho < \alpha) \\ \frac{d^2\omega}{d\rho^2}+\left(\lambda-\gamma-\frac{S''}{4S^2}+\frac{5}{16}\frac{S'^2}{S^3}-\frac{n^2-\frac{1}{4}}{Sr^2}\right)\omega &= 0 \quad (\alpha \leqslant \rho \leqslant \beta) \end{aligned}\right\}, \tag{6.3}$$

where $\rho = \alpha, \beta$ correspond to $r = a, b$.

7. Suppose first that $s(x,y)$ is bounded, say $s(x,y) \leqslant 1$. Then we may take $S(r) = 1$ in the above equations, $\rho = r$, and

$$\left.\begin{aligned} \frac{d^2\omega}{d\rho^2}+\left(\lambda-\frac{n^2-\frac{1}{4}}{\rho^2}\right)\omega &= 0 \quad (0 < \rho < \alpha) \\ \frac{d^2\omega}{d\rho^2}+\left(\lambda-\gamma-\frac{n^2-\frac{1}{4}}{\rho^2}\right)\omega &= 0 \quad (\alpha \leqslant \rho \leqslant \beta) \end{aligned}\right\}. \tag{7.1}$$

If $n^2-\frac{1}{4} > \gamma\alpha^2$, the coefficient of ω in each equation is negative for $\lambda < \gamma$, and so there are no eigenvalues in this range. Otherwise an eigenfunction is (apart from a normalizing factor)

$$\omega(\rho) = \begin{cases} J_n(\rho\sqrt{\lambda}) & (0 < \rho < \alpha), \\ AI_n\{\rho\sqrt{(\gamma-\lambda)}\}+BK_n\{\rho\sqrt{(\gamma-\lambda)}\} & (\alpha \leqslant \rho \leqslant \beta), \end{cases}$$

where, from the continuity of ω and ω' at $\rho = \alpha$,

$$A = \alpha[\sqrt{\lambda}J_n'(\alpha\sqrt{\lambda})K_n\{\alpha\sqrt{(\gamma-\lambda)}\}-\sqrt{(\gamma-\lambda)}J_n(\alpha\sqrt{\lambda})K_n'\{\alpha\sqrt{(\gamma-\lambda)}\}],$$
$$B = -\alpha[\sqrt{\lambda}J_n'(\alpha\sqrt{\lambda})I_n\{\alpha\sqrt{(\gamma-\lambda)}\}-\sqrt{(\gamma-\lambda)}J_n(\alpha\sqrt{\lambda})I_n'\{\alpha\sqrt{(\gamma-\lambda)}\}].$$

The condition $\omega(\beta) = 0$ then gives

$$\frac{\sqrt{\lambda}J_n'(\alpha\sqrt{\lambda})K_n\{\alpha\sqrt{(\gamma-\lambda)}\}-\sqrt{(\gamma-\lambda)}J_n(\alpha\sqrt{\lambda})K_n'\{\alpha\sqrt{(\gamma-\lambda)}\}}{\sqrt{\lambda}J_n'(\alpha\sqrt{\lambda})I_n\{\alpha\sqrt{(\gamma-\lambda)}\}-\sqrt{(\gamma-\lambda)}J_n(\alpha\sqrt{\lambda})I_n'\{\alpha\sqrt{(\gamma-\lambda)}\}} = \frac{K_n\{\beta\sqrt{(\gamma-\lambda)}\}}{I_n\{\beta\sqrt{(\gamma-\lambda)}\}}, \tag{7.2}$$

and the eigenvalues are the roots of these equations for $0 \leqslant n \leqslant \sqrt{(\gamma\alpha^2+\frac{1}{4})}$.

Suppose first that $n > 0$. Then, as $\lambda \to \gamma$, the left-hand side†

$$\sim \frac{1}{2n\alpha^{2n}(\gamma-\lambda)^n}\frac{\alpha\sqrt{\gamma}J_n'(\alpha\sqrt{\gamma})+nJ_n(\alpha\sqrt{\gamma})}{\alpha\sqrt{\gamma}J_n'(\alpha\sqrt{\gamma})-nJ_n(\alpha\sqrt{\gamma})},$$

and we can suppose, by choice of α, that the coefficient of $(\gamma-\lambda)^{-n}$ is not 0 or ∞. Hence there is an interval $(\gamma-\delta, \gamma)$, independent of β, in which the modulus of the left-hand side of (7.2) is greater than $A(\gamma-\lambda)^{-n}$. Also $x^{2n}K_n(x)/I_n(x)$ is bounded for all positive x, and so the right-hand side of (7.2) is less than

$$A\beta^{-2n}(\gamma-\lambda)^{-n}.$$

There are therefore no roots in the interval $(\gamma-\delta, \gamma)$ if β is large enough.

If $n = 0$, the left-hand side of (7.2)

$$\sim \log(\gamma-\lambda)^{-\frac{1}{2}}$$

as $\lambda \to \gamma$. If $\beta\sqrt{(\gamma-\lambda)} \leqslant 1$, the right-hand side is less than

$$\log(\gamma-\lambda)^{-\frac{1}{2}}-\log\beta+O(1).$$

Hence there are no roots in this case if β is large enough. We can also suppose the interval $(\gamma-\delta, \gamma)$ to be chosen so that the modulus of the left-hand side exceeds $\max\limits_{x \geqslant 1} K_0(x)/I_0(x)$, and so there are no roots here either.

It was shown in (4) that the number of roots in any given interval $\lambda_0 < \lambda < \gamma-\delta$ is ultimately independent of β. Thus the total number of eigenvalues of (6.1) less than γ is bounded as $b \to \infty$. This gives the desired result on the discreteness of the spectrum if $s(x, y) \leqslant 1$.

† See Watson, *Theory of Bessel Functions*, § 3.7 (2), § 3.71 (14), (15).

8. Suppose next that $s(x,y) = O(r^k)$ as $r \to \infty$, where k is a given positive number. We may suppose that

$$s(x,y) \leqslant 1 \quad (0 < r \leqslant 1), \qquad s(x,y) \leqslant r^k \quad (r > 1).$$

Let
$$S(r) = 1 \quad (0 < r \leqslant 1), \qquad S(r) = r^k \quad (r > 1).$$

Then

$$\rho = r \quad (0 < r \leqslant 1), \qquad \rho = 1+(\tfrac{1}{2}k+1)^{-1}(r^{\frac{1}{2}k+1}-1) \quad (r > 1).$$

Supposing that $\alpha > 1$, (6.3) then gives

$$\frac{d^2\omega}{d\rho^2}+\left(\lambda-\frac{n^2-\frac{1}{4}}{\rho^2}\right)\omega = 0 \quad (0 < \rho \leqslant 1),$$

$$\frac{d^2\omega}{d\rho^2}+\left(\lambda-\frac{\nu^2-\frac{1}{4}}{(\rho-\rho_0)^2}\right)\omega = 0 \quad (1 < \rho < \alpha),$$

$$\frac{d^2\omega}{d\rho^2}+\left(\lambda-\gamma-\frac{\nu^2-\frac{1}{4}}{(\rho-\rho_0)^2}\right)\omega = 0 \quad (\alpha \leqslant \rho \leqslant \beta),$$

where
$$\nu = \frac{n}{\frac{1}{2}k+1}, \qquad \rho_0 = \frac{k}{k+2}.$$

The discussion of these equations follows the same lines as that of (7.1), and the same result as before is obtained. In fact it seems likely that no restriction on $s(x,y)$ is necessary for the result, but, if there is no restriction, the line of proof given here runs into difficulties.

REFERENCES

1. K. O. Friedrichs, 'Criteria for discrete spectra', *Comm. Pure and Applied Math.* 3 (1950), 439–49.
2. D. B. Sears, 'Note on the uniqueness of the Green's functions associated with certain differential equations', *Canadian J. of Math.* 2 (1950), 314–25.
3. E. C. Titchmarsh, 'On the uniqueness of the Green's function associated with a second-order differential equation,' ibid. 1 (1949), 191–8.
4. —— 'On the discreteness of the spectrum associated with certain differential equations', *Annali di Mat.* (IV) 28 (1949), 141–7.
5 —— 'Eigenfunction expansions associated with partial differential equations', *Proc. London Math. Soc.* (3) 1 (1951), 1–27.
6. —— '—— (II)', ibid. 3 (1953), 80–98.
7. —— '—— (III)', ibid. 153–69.

INDUCTIVE DIMENSION OF COMPLETELY NORMAL SPACES

By C. H. DOWKER (*London*)

[Received 13 March 1953]

USING the dimension defined inductively in terms of closed sets, I show that in a completely normal space the dimension of the union of two disjoint sets, one of which is closed, is at most equal to the greatest of their dimensions. A corresponding theorem is proved for a countable union of disjoint sets. It follows that in completely normal spaces the subset theorem implies the sum theorem. The subset theorem and the open subset theorem are shown to be equivalent. Therefore (Theorem 1) the sum and subset theorems hold for any completely normal space in which the dimension of a set A is never less than the dimension of a relatively open subset of A.

E. Čech (**3**) extended the sum and subset theorems from separable metric spaces to perfectly normal spaces. I introduce a new class of normal spaces, intermediate between completely normal and perfectly normal, which I call *totally normal*. A normal space X is *totally normal* if each open set G of X has a locally finite covering by open subsets each of which is an F_σ set of X. The totally normal spaces include the hereditarily paracompact Hausdorff spaces as well as the perfectly normal spaces. It is shown that the open subset theorem and hence the subset theorem and sum theorem hold for totally normal spaces. The covering theorem of Čech also holds for totally normal spaces.

The open subset theorem does not hold for all normal Hausdorff spaces nor for all completely normal spaces. The question of whether it holds for all completely normal Hausdorff spaces is still undecided.

1. Definitions and known theorems

A space X is called *normal* if for each pair of disjoint closed sets E and F of X there exist disjoint open sets U and V with $E \subset U$ and $F \subset V$. It is no restriction on the space X to require also that the closures $\bar{U}$ and $\bar{V}$ of U and V should be disjoint or that the open sets U and V should be F_σ sets, i.e. countable unions of closed sets. A space X is called *completely normal* if every subset of X is a normal space. Clearly every subset of a completely normal space is completely normal.

[1.1] *If every open subset of a space X is a normal space, X is completely normal.*

Proof. Let every open subset of X be normal and let A be an arbitrary subset of X. Let E and F be any two disjoint sets closed in A. Let $G = X - \bar{E} \cap \bar{F}$, $E_1 = \bar{E} \cap G$, and $F_1 = \bar{F} \cap G$. Then $A \subset G$, G is open and hence normal, and E_1 and F_1 are disjoint closed sets of G. Hence there exist disjoint open sets U_1 and V_1 of G with $E_1 \subset U_1$ and $F_1 \subset V_1$. Let $U = U_1 \cap A$ and $V = V_1 \cap A$; then U and V are open in A, $E \subset U$, $F \subset V$ and $U \cap V = 0$. Therefore A is normal. Therefore X is completely normal.

The inductive dimension of a space X, $\operatorname{Ind} X$, is defined inductively as follows. If X is the empty set, $\operatorname{Ind} X = -1$. For $n = 0, 1,\ldots$, $\operatorname{Ind} X \leqslant n$ means that for every open set G containing E there is an open set U with $E \subset U \subset G$ and $\operatorname{Ind}(\bar{U} - U) \leqslant n-1$. $\operatorname{Ind} X = \infty$ means that there is no n for which $\operatorname{Ind} X \leqslant n$.

[1.2] *If A is any closed subset of a space X,* $\operatorname{Ind} A \leqslant \operatorname{Ind} X$.

Proof.† It is sufficient to show that, if $\operatorname{Ind} X \leqslant n$, then $\operatorname{Ind} A \leqslant n$. This is trivially true for dimension -1 and we assume it true for dimension $n-1$. Let $\operatorname{Ind} X \leqslant n$, let A be closed in X, and let $E \subset G \subset A$ with E closed in A and G open in A. Then E is closed in X and there exists G_1 open in X with $G_1 \cap A = G$, and hence with $E \subset G_1$. Hence there exists U_1 open in X with $E \subset U_1 \subset G_1$ and $\operatorname{Ind}(\bar{U}_1 - U_1) \leqslant n-1$. Let $U = U_1 \cap A$, then U is open in A and $E \subset U \subset G$. Then $\bar{U} \subset \bar{U}_1$, and hence

$$\bar{U} \cap A - U \subset \bar{U}_1 \cap A - U = \bar{U}_1 \cap A - U_1 \cap A = (\bar{U}_1 - U_1) \cap A \subset \bar{U}_1 - U_1.$$

And $\bar{U} \cap A - U$ is closed in A, hence in X and in $\bar{U}_1 - U_1$. Hence, by the induction hypothesis, $\operatorname{Ind}(\bar{U} \cap A - U) \leqslant n-1$. Therefore $\operatorname{Ind} A \leqslant n$, as was to be shown.

[1.3] $\operatorname{Ind} X \leqslant n$ *is equivalent to the following condition on X:*

(α) *If $E \subset G \subset X$ with E closed and G open, then X is the union of three disjoint sets U, V, and C with U and V open, $E \subset U \subset G$ and* $\operatorname{Ind} C \leqslant n-1$.

Proof. If there exists U with $E \subset U \subset G$ and $\operatorname{Ind}(\bar{U} - U) \leqslant n-1$, we set $C = \bar{U} - U$ and $V = X - \bar{U}$. Then U, V, and C are disjoint, $X = U \cup V \cup C$, U and V are open, and $\operatorname{Ind} C \leqslant n-1$.

If, on the other hand, X is the union of disjoint sets U, V, and C with U and V open, $E \subset U \subset G$ and $\operatorname{Ind} C \leqslant n-1$, then U is contained in the closed set $U \cup C$ and hence $\bar{U} - U \subset C$. Therefore, since $\bar{U} - U$ is closed, $\operatorname{Ind}(\bar{U} - U) \leqslant \operatorname{Ind} C \leqslant n-1$.

† See Čech (**3**), proposition 16.2.

[1.4] *If X is normal,* $\operatorname{Ind} X \leqslant n$ *is equivalent to the following condition:*

(β) *If E and F are disjoint closed sets of X, then X is the union of disjoint sets U, V, and C with U and V open, $E \subset U$, $F \subset V$, and*

$$\operatorname{Ind} C \leqslant n-1.$$

Proof.† It is sufficient to show that conditions (α) and (β) are equivalent when X is normal. First let (α) be satisfied and let E and F be disjoint closed sets of X. Since X is normal, there exists an open set G such that $E \subset G \subset \bar{G} \subset X-F$. Then, by ($\alpha$), X is the union of disjoint sets U, V, and C with U and V open, $E \subset U \subset G$, and $\operatorname{Ind} C \leqslant n-1$. Since $U \subset G$, $\bar{U} \subset \bar{G} \subset X-F$ and hence $F \subset X-\bar{U} = V$. Thus (β) is satisfied.

Conversely let (β) be satisfied and let $E \subset G \subset X$ with E closed and G open. Let $F = X-G$; then E and F are closed and disjoint. Hence, by (β), X is the union of disjoint sets U, V, and C with U and V open, $E \subset U$, $F \subset V$, and $\operatorname{Ind} C \leqslant n-1$. Since $X-G = F \subset V \subset X-U$, therefore $U \subset G$. Thus (α) is satisfied.

2. Sum theorem for disjoint sets

In the following lemma I show that, if a completely normal space Y is the union of a sequence $\{D_i\}$ of disjoint sets such that the partial unions $\bigcup_{j \leqslant i} D_j$ are closed in Y, then $\operatorname{Ind} Y \leqslant \sup \operatorname{Ind} D_i$. In particular, if $Y = D_1 \cup D_2$ with $D_1 \cap D_2 = 0$ and D_1 closed, then

$$\operatorname{Ind} Y \leqslant \max(\operatorname{Ind} D_1, \operatorname{Ind} D_2).$$

This lemma does not extend to arbitrary normal spaces. O. V. Lokutzievski (**5**) has given an example of a space S and a closed subset S_1 such that S, S_1, and $S-S_1$ are normal, $\operatorname{Ind} S_1 = 1$, $\operatorname{Ind}(S-S_1) = 1$ but $\operatorname{Ind} S = 2$.

[2.1] *Let Y_i $(i = 1, 2,\ldots)$ be open sets in a completely normal space Y such that $Y = Y_1 \supset Y_2 \supset \ldots$, $\bigcap_{i=1}^{\infty} Y_i = 0$ and, for each i,* $\operatorname{Ind}(Y_i - Y_{i+1}) \leqslant n$. *Then* $\operatorname{Ind} Y \leqslant n$.

Proof. This is trivially true for dimension -1 and we assume it true for dimension $n-1$. Let E and F be any two disjoint closed sets of Y. Since Y is normal, there exist open sets U_0 and V_0 with $E \subset U_0$, $F \subset V_0$, and $\bar{U}_0 \cap \bar{V}_0 = 0$.

† See Čech (**3**), § 18. Condition (β) is closely related to the original definition of dimension by L. E. J. Brouwer (**2**).

Let $D_i = Y_i - Y_{i+1}$. We construct disjoint sets U_i, V_i, and C_i for $i = 1, 2, \ldots$, with U_i and V_i open in Y_i and hence open in Y and with $C_i \subset D_i \subset U_i \cup V_i \cup C_i$, $\operatorname{Ind} C_i \leqslant n-1$, $\overline{U}_i \cap \overline{V}_i \cap Y_{i+1} = 0$, $U_i \supset \overline{U}_{i-1} \cap Y_i$, and $V_i \supset \overline{V}_{i-1} \cap Y_i$.

We already have the sets U_0 and V_0, and, assuming U_{i-1} and V_{i-1} have been constructed so that $\overline{U}_{i-1} \cap \overline{V}_{i-1} \cap Y_i = 0$, we construct the sets U_i, V_i, and C_i as follows.

The sets $\overline{U}_{i-1} \cap D_i$ and $\overline{V}_{i-1} \cap D_i$ are disjoint and closed in D_i and $\operatorname{Ind} D_i \leqslant n$. Hence D_i is the union of disjoint sets G_i, H_i, and C_i, with G_i and H_i open in D_i and with $\overline{U}_{i-1} \cap D_i \subset G_i$, $\overline{V}_{i-1} \cap D_i \subset H_i$ and $\operatorname{Ind} C_i \leqslant n-1$. Then C_i is closed in $D_i = Y_i - Y_{i+1}$ which is closed in Y_i; hence $Y_i - C_i$ is open in Y_i and hence open in Y.

The sets G_i and H_i are closed in $D_i - C_i$ and hence closed in $Y_i - C_i$. Let $E_i = (\overline{U}_{i-1} \cup G_i) \cap (Y_i - C_i)$ and $F_i = (\overline{V}_{i-1} \cup H_i) \cap (Y_i - C_i)$, then E_i and F_i are closed sets of $Y_i - C_i$. Since $\overline{V}_{i-1} \cap D_i \subset H_i$ and $G_i \subset D_i - H_i$, therefore $\overline{V}_{i-1} \cap G_i = 0$, and similarly $\overline{U}_{i-1} \cap H_i = 0$. Hence, since

$$\overline{U}_{i-1} \cap \overline{V}_{i-1} \cap (Y_i - C_i) \subset \overline{U}_{i-1} \cap \overline{V}_{i-1} \cap Y_i = 0 \quad \text{and} \quad G_i \cap H_i = 0,$$

E_i and F_i are disjoint.

Since Y is completely normal, $Y_i - C_i$ is normal. Hence there exist sets U_i and V_i open in $Y_i - C_i$ and hence open in Y such that $E_i \subset U_i$, $F_i \subset V_i$, and $U_i \cap V_i = 0$ and such that, moreover, $\overline{U}_i \cap \overline{V}_i \cap (Y_i - C_i) = 0$. Then $\overline{U}_i \cap \overline{V}_i \cap Y_{i+1} = 0$, and, since U_i and V_i are disjoint and contained in $Y_i - C_i$, the sets U_i, V_i, and C_i are disjoint.

Since $D_i = G_i \cup H_i \cup C_i$ and $G_i \subset E_i \subset U_i$ and $H_i \subset F_i \subset V_i$, therefore $C_i \subset D_i \subset U_i \cup V_i \cup C_i$. Since $\overline{U}_{i-1} \cap D_i \subset G_i$, therefore $\overline{U}_{i-1} \cap C_i = 0$ and hence $\overline{U}_{i-1} \cap Y_i = \overline{U}_{i-1} \cap (Y_i - C_i) \subset E_i \subset U_i$. Similarly $V_i \supset \overline{V}_{i-1} \cap Y_i$. Thus the sets U_i, V_i, and C_i have the required properties.

Let $U = \bigcup_{i=0}^{\infty} U_i$, $V = \bigcup_{i=0}^{\infty} V_i$, $Z_i = \bigcup_{j=i}^{\infty} C_j$, and $C = Z_1 = \bigcup_{j=1}^{\infty} C_j$. Then the sets U and V are unions of open sets and hence open; and

$$E \subset U_0 \subset U \quad \text{and} \quad F \subset V_0 \subset V.$$

Every point of Y is in some D_i, hence in U_i, V_i, or C_i and hence in U, V, or C; thus $Y \subset U \cup V \cup C$.

If $i \leqslant j$, $U_i \cap Y_j \subset U_j$ and $V_i \cap Y_j \subset V_j$. Therefore $U_i \cap V_j \subset U_j \cap V_j = 0$ and $U_j \cap V_i \subset U_j \cap V_j = 0$, and hence $U \cap V = 0$. If $i \leqslant j$,

$$U_i \cap C_j \subset U_j \cap C_j = 0$$

and, if $i > j$, since $U_i \subset Y_i$, $U_i \cap C_j \subset Y_i \cap C_j = 0$. Hence $U \cap C = 0$ and similarly $V \cap C = 0$. Thus the sets U, V, and C are disjoint.

As a subset of a completely normal space, C is completely normal. Each $Z_i = C \cap Y_i$ is open in C, $Z_i \supset Z_{i-1}$ and

$$\bigcap_{i=1}^{\infty} Z_i \subset \bigcap_{i=1}^{\infty} Y_i = 0.$$

We have $Z_i = \bigcup_{j=i}^{\infty} C_j = C_i \cup Z_{i+1}$ and, for $i < j$, $C_i \cap C_j \subset C_i \cap Y_j = 0$ and hence $C_i \cap Z_{i+1} = 0$. Therefore $C_i = Z_i - Z_{i+1}$ and

$$\operatorname{Ind}(Z_i - Z_{i+1}) = \operatorname{Ind} C_i \leqslant n-1.$$

Therefore, by the induction hypothesis, $\operatorname{Ind} C \leqslant n-1$. Thus condition (β) is satisfied and $\operatorname{Ind} Y \leqslant n$, as was to be shown.

[2.2] *If A is a closed subset of a completely normal space Y and if* $\operatorname{Ind} A \leqslant n$ *and* $\operatorname{Ind}(Y-A) \leqslant n$, *then* $\operatorname{Ind} Y \leqslant n$.

Proof. Let $Y_1 = Y$, $Y_2 = Y-A$, and $Y_3 = Y_4 = \ldots = 0$. Then

$$Y = Y_1 \supset Y_2 \supset Y_3 \supset \ldots, \qquad \bigcap_{i=1}^{\infty} Y_i = 0,$$

and $\operatorname{Ind}(Y_1 - Y_2) = \operatorname{Ind} A \leqslant n$, $\operatorname{Ind}(Y_2 - Y_3) = \operatorname{Ind}(Y-A) \leqslant n$, and, for $i \geqslant 3$, $\operatorname{Ind}(Y_i - Y_{i+1}) = -1 \leqslant n$. Therefore, by [2.1], $\operatorname{Ind} Y \leqslant n$.

3. Consequences of the open subset theorem

In order to discuss the relations between the subset theorem, the open subset theorem, and the sum theorem, we consider the following conditions which a space X may satisfy.

(a_n) If $B \subset A \subset X$ and $\operatorname{Ind} A \leqslant n$, then $\operatorname{Ind} B \leqslant n$.

(b_n) If $G \subset A \subset X$ with G open in A and $\operatorname{Ind} A \leqslant n$, then $\operatorname{Ind} G \leqslant n$.

(c_n) If $A = B \cup C \subset X$ with B closed in A, $\operatorname{Ind} B \leqslant n$ and $\operatorname{Ind} C \leqslant n$, then $\operatorname{Ind} A \leqslant n$.

(d_n) If $A = \bigcup_{i=1}^{\infty} A_i \subset X$ with each A_i closed in A and $\operatorname{Ind} A_i \leqslant n$, then $\operatorname{Ind} A \leqslant n$.

If $Y \subset X$ and X satisfies condition (a_n) [or (b_n), (c_n), (d_n)], then Y also satisfies (a_n) [or (b_n), (c_n), (d_n)].

Clearly (a_n) implies (b_n).

[3.1] *If a space X satisfies (a_{n-1}) and (b_n), then it satisfies (a_n).*

Proof. Let X satisfy conditions (a_{n-1}) and (b_n), let $B \subset A \subset X$, and let $\operatorname{Ind} A \leqslant n$. Let $E \subset G \subset B$ with E closed in B and G open in B. Then there exist E_1 closed in A and G_1 open in A with $E_1 \cap B = E$ and $G_1 \cap B = G$. Let $H = (A - E_1) \cup G_1$, then $H \supset (B-E) \cup G = B$. By

(b_n), since H is open in A, $\operatorname{Ind} H \leqslant n$. We have $E_1 \cap H$ closed in H, G_1 open in H, and

$$E_1 \cap H = E_1 \cap [(A-E_1) \cup G_1] = E_1 \cap G_1 \subset G_1.$$

Therefore, since $\operatorname{Ind} H \leqslant n$, there exists U_1 open in H with

$$E_1 \cap H \subset U_1 \subset G_1 \quad \text{and} \quad \operatorname{Ind}(\overline{U}_1 \cap H - U_1) \leqslant n-1.$$

Let $U = U_1 \cap B$. Then $E = E_1 \cap B \subset E_1 \cap H \subset U_1$ and $E \subset B$; hence $E \subset U$. And $U = U_1 \cap B \subset G_1 \cap B = G$; thus $E \subset U \subset G$. The boundary of U in B is

$$\overline{U} \cap B - U = \overline{U} \cap B - U_1 \cap B \subset \overline{U}_1 \cap B - U_1 \cap B = (\overline{U}_1 - U_1) \cap B,$$

and, since $B \subset H$,

$$(\overline{U}_1 - U_1) \cap B = (\overline{U}_1 \cap H - U_1) \cap B \subset \overline{U}_1 \cap H - U_1.$$

Therefore, by (a_{n-1}), $\operatorname{Ind}(\overline{U} \cap B - U) \leqslant n-1$. Thus $\operatorname{Ind} B \leqslant n$, and condition (a_n) is satisfied. This completes the proof.

Thus we have the implications: $(a_{n-1})+(b_n) \to (a_n) \to (b_n)$. The condition (a_{-1}) is trivially satisfied. Hence, if (b_n) is satisfied for every n, (a_n) is also satisfied for every n. Thus the open subset theorem implies the subset theorem.

[3.2] *If X is completely normal, (b_n) implies (c_n).*

Proof. Let X be a completely normal space satisfying (b_n) and let $A = B \cup C \subset X$ with B closed in A, and with $\operatorname{Ind} B \leqslant n$ and $\operatorname{Ind} C \leqslant n$. Since B is closed in A, $A-B$ is open in A, hence open in C, and hence, by (b_n), $\operatorname{Ind}(A-B) \leqslant n$. Since X is completely normal, so is A; and B is closed in A, $\operatorname{Ind} B \leqslant n$ and $\operatorname{Ind}(A-B) \leqslant n$. Hence, by [2.2], $\operatorname{Ind} A \leqslant n$. Thus condition (c_n) is satisfied.

[3.3] *If X is completely normal, (b_n) implies (d_n).*

Proof. Let X be a completely normal space satisfying (b_n) and let $A = \bigcup_{i=1}^{\infty} A_i \subset X$ with A_i closed in A and $\operatorname{Ind} A_i \leqslant n$. Let $D_i = A_i - \bigcup_{j<i} A_j$;

then
$$A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} D_i.$$

Let
$$Y_i = \bigcup_{j=i}^{\infty} D_j = A - \bigcup_{j<i} A_j;$$

then $Y_i \supset Y_{i+1}$ and
$$\bigcap_{i=1}^{\infty} Y_i = A - \bigcup_{j=1}^{\infty} A_j = 0.$$

Since $\bigcup_{j<i} A_j$ is closed in A, D_i is open in A_i and Y_i is open in A. Then, by (b_n), $\operatorname{Ind} D_i \leqslant n$. The sets D_i are disjoint and

$$Y_i = \bigcup_{j=i}^{\infty} D_j = D_i \cup Y_{i+1}.$$

Hence $D_i = Y_i - Y_{i+1}$. Therefore $\operatorname{Ind}(Y_i - Y_{i+1}) \leqslant n$. Since $A \subset X$, A is completely normal. Therefore, by [2.1], $\operatorname{Ind} A \leqslant n$. Thus condition (d_n) is satisfied.

Thus the open subset theorem implies the sum theorem for completely normal spaces.

THEOREM 1. *If X is a completely normal space satisfying condition (b_n) for all n, then X also satisfies (a_n), (c_n), and (d_n) for all n.*

Proof. This follows from [3.1], [3.2], and [3.3].

4. Totally normal spaces

I introduce a class of normal spaces for which the subset theorem and sum theorem of inductive dimension can be shown to hold.

Definition. A space X is called *totally normal* if it is normal and if each open set G of X is the union of a collection $\{G_\alpha\}$, locally finite in G, of open F_σ sets of X.

It will be shown that perfectly normal spaces are totally normal and totally normal spaces are completely normal.

[4.1] *Perfectly normal spaces are totally normal.*

Proof. Each open set G of a perfectly normal space X is an F_σ set of X. Thus the collection $\{G_\alpha\}$ of open F_σ sets covering G may consist of the one set G.

[4.2] *Hereditarily paracompact Hausdorff spaces are totally normal.*

Proof. Let X be an hereditarily paracompact Hausdorff space and let G be an open set in X. Then X and G are normal Hausdorff spaces (4). Then X is regular and hence each point x of G is contained in an open set U_x whose closure $\overline{U}_x$ is contained in G. The sets U_x form a covering of the paracompact space G; hence there is a locally finite refinement $\{V_\alpha\}$ of the covering by the sets U_x. Then (4) there is a covering $\{F_\alpha\}$ of G by closed sets of G such that $F_\alpha \subset V_\alpha$. Hence, since G is normal, there exist open F_σ sets G_α of G with $F_\alpha \subset G_\alpha \subset V_\alpha$. Since $G = \bigcup F_\alpha$ and $F_\alpha \subset G_\alpha \subset G$, therefore $G = \bigcup G_\alpha$. Since $G_\alpha \subset V_\alpha$ and $\{V_\alpha\}$ is locally finite in G, therefore $\{G_\alpha\}$ is locally finite in G. Since G_α is open in the open set G, it is open in X. Since G_α is an F_σ set in G and, for some x, $G_\alpha \subset \overline{V}_\alpha \subset \overline{U}_x \subset G$,

therefore G_α is an F_σ set of the closed set $\overline{V}_\alpha$ of X and hence is an F_σ set of X. Thus the space X is totally normal.

Let us consider the following two conditions on an open set G of a space X.

(h) G is the union of a collection, locally finite in G, of open F_σ sets of X.

(k) For each $i = 1, 2, \ldots$, there is a collection $\{W_{i\alpha}\}$, locally finite in G, of disjoint open sets and a corresponding collection $\{F_{i\alpha}\}$ of closed sets of X such that $F_{i\alpha} \subset W_{i\alpha} \subset G$ and such that $\bigcup_{i=1}^{\infty} \bigcup_{\alpha} F_{i\alpha} = G$.

The condition (h) is satisfied by open sets of totally normal spaces.

[4.3] *For an open set G of a normal space X the conditions* (h) *and* (k) *are equivalent.*

Proof. (h) $\to$ (k). Let G be an open set of a normal space X and let $G = \bigcup G_\alpha$, where $\{G_\alpha\}$ is locally finite in G and each G_α is an open F_σ set of X. Then there is a continuous function f_α $(0 \leqslant f_\alpha(x) \leqslant 1)$ such that $f_\alpha(x) > 0$ if and only if $x \in G_\alpha$.

Let the indices α be well ordered and let

$$W_{i\alpha} = \{x \mid f_\alpha(x) > (i+1)^{-1}, f_\beta(x) < (i+1)^{-1} \text{ for all } \beta < \alpha\}.$$

Then from the local finiteness of $\{G_\alpha\}$ in G it follows that $W_{i\alpha}$ is open in G and hence in X. Since $W_{i\alpha} \subset G_\alpha$ and $\{G_\alpha\}$ is locally finite in G, therefore, for each i, $\{W_{i\alpha}\}$ is locally finite in G. Clearly, if $\beta < \alpha$, then

$$W_{i\alpha} \cap W_{i\beta} = 0.$$

Let $F_{i\alpha} = \{x \mid f_\alpha(x) \geqslant 1/i,\ x \notin G_\beta \text{ for } \beta < \alpha\}$. Then $F_{i\alpha}$ is closed in X and $F_{i\alpha} \subset W_{i\alpha} \subset G$. If $x \in G$, let G_α be the first set of the covering $\{G_\alpha\}$ of G which contains x. Then $f_\alpha(x) > 0$ and hence, for some $i, f_\alpha(x) \geqslant 1/i$ while $x \notin G_\beta$ for $\beta < \alpha$; hence $x \in F_{i\alpha}$. Thus $\bigcup_{i=1}^{\infty} \bigcup_{\alpha} F_{i\alpha} = G$. Thus condition ($k$) is satisfied.

(k) $\to$ (h). Let G be an open set of a normal space X and, for $i = 1, 2, \ldots$, let $\{W_{i\alpha}\}$ be a collection, locally finite in G, of disjoint open sets. Let $F_{i\alpha}$ be closed in X, $F_{i\alpha} \subset W_{i\alpha} \subset G$ and let $\bigcup_{i=1}^{\infty} \bigcup_{\alpha} F_{i\alpha} = G$. Since X is normal, there is a continuous function $g_{i\alpha}$ $(0 \leqslant g_{i\alpha}(x) \leqslant 1)$ such that $g_{i\alpha}(x) = 0$ for $x \in X - W_{i\alpha}$ and $g_{i\alpha}(x) = 1$ for $x \in F_{i\alpha}$. Let

$$g_i(x) = \sum_\alpha g_{i\alpha}(x) = \sup_\alpha g_{i\alpha}(x).$$

Then, since $\{W_{i\alpha}\}$ is locally finite in G, $g_i(x)$ exists and is continuous in G. Let

$$G_{i\alpha} = \{x \mid g_{i\alpha}(x) > 0,\ g_j(x) < 1/i \text{ for } j < i\},$$

and let $H_{i\alpha} = \{x \mid g_{i\alpha}(x) > 0\}$. Then $H_{i\alpha}$ is an open F_σ set of X, $H_{i\alpha} \subset G$, and $G_{i\alpha}$ is an open F_σ set in $H_{i\alpha}$ and hence in X.

If $x \in G$, then, for some j and β, $x \in F_{j\beta}$ and hence $g_{j\beta}(x) = 1$ and $g_j(x) = 1$. If i is the least number such that, for some α, $g_{i\alpha}(x) > 0$, then, for $j < i$, $g_j(x) = 0$ and hence $x \in G_{i\alpha}$. Thus $G \subset \bigcup_{i,\alpha} G_{i\alpha}$. Hence, since $G_{i\alpha} \subset H_{i\alpha} \subset G$, $G = \bigcup_{i,\alpha} G_{i\alpha}$.

Since $g_{j\beta}(x) = 1$, there is a neighbourhood N of x in G such that $g_j(y) > \frac{1}{2}$ for $y \in N$. Hence, for $i > j$ and $i > 2$, $N \cap G_{i\alpha} = 0$. Thus N can meet $G_{i\alpha}$ only if $i \leqslant i_0 = \max(2, j)$. For each i, $\{W_{i\alpha}\}$ is locally finite in G and hence some neighbourhood N_i of x meets only a finite number of the sets $W_{i\alpha}$. Since $G_{i\alpha} \subset H_{i\alpha} \subset W_{i\alpha}$, $N_i \cap G_{i\alpha} \neq 0$ for at most a finite number of values of α. Then the intersection $N \cap \bigcap_{i \leqslant i_0} N_i$ is a neighbourhood of x which meets only a finite number of the sets $G_{i\alpha}$. Thus the collection $\{G_{i\alpha}\}$, for all i and all α, is locally finite in G. Thus condition (h) is satisfied.

[4.4] *Let a space X be the union of disjoint sets G_α each of which is open and closed in X. If each G_α is normal, then X is normal.*

Proof. Let E and F be closed and disjoint in X. Then $E \cap G_\alpha$ and $F \cap G_\alpha$ are closed and disjoint in the normal space G_α. Hence there exist disjoint sets U_α and V_α open in G_α, and hence open in X, such that $E \cap G_\alpha \subset U_\alpha$ and $F \cap G_\alpha \subset V_\alpha$. Let $U = \bigcup U_\alpha$ and $V = \bigcup V_\alpha$, then U and V are open and disjoint and $E = \bigcup E \cap G_\alpha \subset \bigcup U_\alpha = U$ and $F \subset V$. Therefore X is normal.

[4.5] *Let X be a space and let $\{C_i\}$ be a sequence of closed sets whose interiors cover X. If each C_i is normal, then X is normal.*

Proof. Let E_0 and F_0 be any two disjoint closed sets of X. We construct increasing sequences $\{E_i\}$ and $\{F_i\}$ of disjoint closed sets of X such that, for $i = 1, 2, \ldots$,

$$E_{i-1} \cap C_i \subset G_i \subset E_i \quad \text{and} \quad F_{i-1} \cap C_i \subset H_i \subset F_i,$$

where G_i and H_i are open and disjoint in C_i.

Assume that we already have E_{i-1} and F_{i-1}. Then $E_{i-1} \cap C_i$ and $F_{i-1} \cap C_i$ are disjoint closed sets of the normal space C_i. Hence there exist disjoint open sets G_i and H_i of C_i such that $E_{i-1} \cap C_i \subset G_i$, $F_{i-1} \cap C_i \subset H_i$, and $\bar{G}_i \cap \bar{H}_i = 0$. Let $E_i = E_{i-1} \cup \bar{G}_i$ and $F_i = F_{i-1} \cup \bar{H}_i$. Then E_i and F_i are disjoint closed sets of X and $G_i \subset E_i$ and $H_i \subset F_i$.

Let $\operatorname{int} C_i$ be the interior of C_i and let

$$U = \bigcup_{i=1}^{\infty} G_i \cap \operatorname{int} C_i \quad \text{and} \quad V = \bigcup_{i=1}^{\infty} H_i \cap \operatorname{int} C_i.$$

Since G_i is open in C_i, $G_i \cap \operatorname{int} C_i$ is open in $\operatorname{int} C_i$, and hence is open in X. Therefore U is open in X and similarly V is open in X. Since

$$G_i \cap \operatorname{int} C_i \subset \bar{G}_i \subset E_i,$$

therefore $U \subset \bigcup E_i$ and similarly $V \subset \bigcup F_i$. And, if $k = \max(i,j)$,

$$E_i \cap F_j \subset E_k \cap F_k = 0.$$

Hence $U \cap V = 0$. Since $X = \bigcup_{i=1}^{\infty} \operatorname{int} C_i$,

$$E_0 = \bigcup_{i=1}^{\infty} E_0 \cap \operatorname{int} C_i \subset \bigcup_{i=1}^{\infty} E_{i-1} \cap \operatorname{int} C_i \subset \bigcup_{i=1}^{\infty} G_i \cap \operatorname{int} C_i = U.$$

Similarly $F_0 \subset V$. Therefore X is normal.

[4.6] *Totally normal spaces are completely normal.*

Proof. Let G be any open set in a totally normal space X. Then, by [4.3], for each $i = 1, 2, \ldots$, there is a collection $\{W_{i\alpha}\}$, locally finite in G, of disjoint open sets and a corresponding collection $\{F_{i\alpha}\}$ of closed sets of X with $F_{i\alpha} \subset W_{i\alpha} \subset G$ and $\bigcup_{i=1}^{\infty} \bigcup_{\alpha} F_{i\alpha} = G$. Then, since X is normal, there exists a closed set $C_{i\alpha}$ with $F_{i\alpha} \subset \operatorname{int} C_{i\alpha} \subset C_{i\alpha} \subset W_{i\alpha}$. Since X is normal and $C_{i\alpha}$ is closed in X, $C_{i\alpha}$ is normal. Let $C_i = \bigcup_{\alpha} C_{i\alpha}$. Since $C_{i\alpha} \subset W_{i\alpha}$, then, for each i, $\{C_{i\alpha}\}$ is a collection of disjoint sets locally finite in G and hence locally finite in C_i. Hence each $C_{i\alpha}$ is open as well as closed in C_i. Hence, by [4.4], C_i is normal. Also $F_{i\alpha} \subset \operatorname{int} C_{i\alpha} \subset \operatorname{int} C_i$ and hence $G = \bigcup_{i,\alpha} F_{i\alpha} \subset \bigcup_{i} \operatorname{int} C_i$. Hence, by [4.5], G is normal. Thus every open set of X is normal and hence, by [1.1], X is completely normal.

[4.7] *Every subset of a totally normal space is totally normal.*

Proof. Let $A \subset X$ with X totally normal. Then, by [4.6], A is normal. Let G be any open set in A. Then there is an open set H of X such that $H \cap A = G$. Since X is totally normal, H is the union of a collection $\{H_\alpha\}$, locally finite in H, of open F_σ sets of X. Then, if $G_\alpha = H_\alpha \cap A$,

$$G = H \cap A = \bigcup H_\alpha \cap A = \bigcup G_\alpha,$$

and G_α is an open F_σ set of A. Each point x of $G \subset H$ has a neighbourhood N in H which meets only a finite number of the sets H_α and hence has the neighbourhood $N \cap G$ in G which meets only a finite number of the sets G_α. Hence $\{G_\alpha\}$ is locally finite in G. Therefore A is totally normal.

Examples. Let Z be a space consisting of a non-countable set of points one of which is distinguished and called z_0. A subset of Z is called *open* (i) if it does not contain z_0 or (ii) if its complement is finite. Then Z is an hereditarily paracompact Hausdorff space which is not perfectly normal. Bing's example H [(**1**), 185] is perfectly normal but not paracompact and so not hereditarily paracompact. His example G [(**1**), 184] is totally normal but neither perfectly normal nor paracompact. The space consisting of the ordinal numbers $\leqslant \omega_1$ with the usual topology is completely normal but not totally normal.

5. Dimension in totally normal spaces

Returning to the theory of inductive dimension, I show (Theorem 2) that the subset theorem, and consequently the sum theorem, holds for totally normal spaces.

[5.1] *Let a space be the union of disjoint sets G_α each of which is open and closed in X. If each* $\operatorname{Ind} G_\alpha \leqslant n$, *then* $\operatorname{Ind} X \leqslant n$.

Proof. This is trivially true for dimension -1, and we assume it true for dimension $n-1$. Let $E \subset W \subset X$ with E closed and W open. Then $E \cap G_\alpha \subset W \cap G_\alpha \subset G_\alpha$ with $E \cap G_\alpha$ closed in G_α and $W \cap G_\alpha$ open in G_α. Hence, since $\operatorname{Ind} G_\alpha \leqslant n$, G_α is the union of disjoint sets U_α, V_α, and C_α with U_α and V_α open in G_α and hence open in X, $E \cap G_\alpha \subset U_\alpha \subset W \cap G_\alpha$ and $\operatorname{Ind} C_\alpha \leqslant n-1$. Let

$$U = \bigcup U_\alpha, \quad V = \bigcup V_\alpha, \quad \text{and} \quad C = \bigcup C_\alpha.$$

Then U, V, and C are disjoint, their union is X, and U and V are open sets. Also

$$E = \bigcup_\alpha E \cap G_\alpha \subset \bigcup_\alpha U_\alpha = U$$

and

$$U = \bigcup_\alpha U_\alpha \subset \bigcup_\alpha W \cap G_\alpha = W.$$

Each set $C_\alpha = C \cap G_\alpha$ is open and closed in C and $\operatorname{Ind} C_\alpha \leqslant n-1$; hence, by the induction hypothesis, $\operatorname{Ind} C \leqslant n-1$. Thus $\operatorname{Ind} X \leqslant n$, as was to be shown.

[5.2] *Let X be a normal space satisfying the condition (d_{n-1}) of* § 3. *Let $\{C_i\}$ and $\{F_i\}$ be sequences of closed sets of X such that $F_i \subset \operatorname{int} C_i$, $X = \bigcup_{i=1}^{\infty} F_i$, and* $\operatorname{Ind} C_i \leqslant n$. *Then* $\operatorname{Ind} X \leqslant n$.

Proof. Let $E \subset G \subset X$ with E closed and G open. Then, since X is normal, there is a closed set K and a sequence $\{W_i\}$ of open sets such that

$$E \subset K \subset \overline{W}_{i+1} \subset W_i \subset G$$

and

$$K = \bigcap_{i=1}^{\infty} W_i.$$

Then $F_i \cap K \subset (\operatorname{int} C_i) \cap W_i$, $F_i \cap K$ is closed and $(\operatorname{int} C_i) \cap W_i$ is open and is contained in C_i. Hence, since $\operatorname{Ind} C_i \leqslant n$, there exists U_i open in C_i with $F_i \cap K \subset U_i \subset (\operatorname{int} C_i) \cap W_i$ and $\operatorname{Ind}(\overline{U}_i \cap C_i - U_i) \leqslant n-1$. Since U_i is open in C_i and $U_i \subset \operatorname{int} C_i$, therefore U_i is open in $\operatorname{int} C_i$ and hence open in X. And, since $U_i \subset C_i$ and C_i is closed, $\overline{U}_i \cap C_i = \overline{U}_i$; hence

$$\operatorname{Ind}(\overline{U}_i - U_i) \leqslant n-1.$$

Let $U = \bigcup_{i=1}^{\infty} U_i$, then U is open in X and

$$E \subset K = \bigcup_i (F_i \cap K) \subset \bigcup_i U_i = U$$

and

$$U \subset \bigcup_i W_i \subset G;$$

thus $E \subset U \subset G$.

Let $x \notin K$; then, for some j, $x \notin \overline{W}_j$, and hence x has a neighbourhood $X - \overline{W}_j$ which meets at most a finite number of the sets U_i $(\subset W_i)$ and hence $x \in \overline{U}$ if and only if $x \in \overline{U}_i$ for some i. Hence, since $K \subset U \subset \overline{U}$ and $K \subset \bigcup U_i \subset \bigcup \overline{U}_i$, we have $\overline{U} = \bigcup_{i=1}^{\infty} \overline{U}_i$. Hence

$$\overline{U} - U = \bigcup \overline{U}_i - \bigcup U_i \subset \bigcup (\overline{U}_i - U_i).$$

Since U_i is open, $\overline{U}_i - U_i$ is closed, and we have $\operatorname{Ind}(\overline{U}_i - U_i) \leqslant n-1$; hence, by (d_{n-1}), $\operatorname{Ind}[\bigcup (\overline{U}_i - U_i)] \leqslant n-1$. Hence, since $\overline{U} - U$ is a closed subset of $\bigcup (\overline{U}_i - U_i)$, $\operatorname{Ind}(\overline{U} - U) \leqslant n-1$. Thus $\operatorname{Ind} X \leqslant n$, as was to be shown.

[5.3] *If X is a totally normal space, condition (d_{n-1}) implies (b_n).*

Proof. Let X be a totally normal space satisfying condition (d_{n-1}) and let $G \subset A \subset X$ with G open in A and $\operatorname{Ind} A \leqslant n$. By [4.7], A is totally normal and hence, by [4.3], there is, for each $i = 1, 2, \ldots$, a collection $\{W_{i\alpha}\}$, locally finite in G, of disjoint open sets of A and a collection $\{F_{i\alpha}\}$ of closed sets of A with

$$F_{i\alpha} \subset W_{i\alpha} \subset G$$

and

$$\bigcup_{i=1}^{\infty} \bigcup_{\alpha} F_{i\alpha} = G.$$

Since A is normal, there exist $V_{i\alpha}$ open in A and $C_{i\alpha}$ closed in A with $F_{i\alpha} \subset V_{i\alpha} \subset C_{i\alpha} \subset W_{i\alpha}$. Then, for each i, the sets $C_{i\alpha}$ are disjoint, $\{C_{i\alpha}\}$ is locally finite in G and hence, if $C_i = \bigcup_{\alpha} C_{i\alpha}$, $\{C_{i\alpha}\}$ is a locally finite collection of disjoint closed sets of C_i. Therefore $C_{i\alpha}$ is open and closed in C_i. Since $C_{i\alpha}$ is closed in A, $\operatorname{Ind} C_{i\alpha} \leqslant n$. Hence, by [5.1], $\operatorname{Ind} C_i \leqslant n$.

Let $F_i = \bigcup_{\alpha} F_{i\alpha}$ and $V_i = \bigcup_{\alpha} V_{i\alpha}$ as well as $C_i = \bigcup_{\alpha} C_{i\alpha}$. Then F_i and C_i are unions of locally finite collections of closed sets of G, and hence are closed in G, while V_i is a union of open sets of G and hence is open in G. Clearly $F_i \subset V_i \subset C_i$; hence F_i is contained in the interior of C_i with respect to G.

Since X is totally normal, G is normal, and, since X satisfies condition (d_{n-1}), so does G. Hence, by [5.2], since

$$G = \bigcup_{i,\alpha} F_{i\alpha} = \bigcup_{i=1}^{\infty} F_i,$$

$\operatorname{Ind} G \leqslant n$. Thus X satisfies condition (b_n). This completes the proof.

THEOREM 2. *Let $A \subset X$ with X totally normal and* $\operatorname{Ind} X \leqslant n$. *Then* $\operatorname{Ind} A \leqslant n$.

THEOREM 3. *Let a totally normal space X be the union of two sets A and B with A closed and* $\operatorname{Ind} A \leqslant n$ *and* $\operatorname{Ind} B \leqslant n$. *Then* $\operatorname{Ind} X \leqslant n$.

THEOREM 4. *Let $\{A_i\}$ be a sequence of closed sets in a totally normal space and let each* $\operatorname{Ind} A_i \leqslant n$. *Then* $\operatorname{Ind} \bigcup_{i=1}^{\infty} A_i \leqslant n$.

Proof. Let X be a totally normal space. Then X is completely normal and hence, by [3.3], (b_n) implies (d_n), and, by [5.3], (d_{n-1}) implies (b_n). Hence, since (b_{-1}) and (d_{-1}) are trivially satisfied, (b_n) and (d_n) hold for all n. Hence, by Theorem 1, (a_n) and (c_n) also hold for all n. Then Theorems 2, 3, 4 follow respectively from (a_n), (c_n), (d_n).

Examples. Let I be the segment $0 \leqslant x \leqslant 1$, and let J be a space consisting of the points of I and a special point j_0; a set of J is open (i) if it is the whole space J or (ii) if it is an open set of I. The space J is trivially normal; there are no disjoint non-empty closed sets. Any subset either is a subset of I and hence is normal or it contains j_0 and hence is trivially normal. Thus J is completely normal. If A is non-empty and closed in J, then $j_0 \in A$, the only open set containing A is J, and $\bar{J} - J = 0$. Hence $\operatorname{Ind} J = 0$. But I is an open set in J and $\operatorname{Ind} I = 1$. Thus the open subset theorem does not hold in the completely normal space J.

Of course J is not a Hausdorff space. An example is given elsewhere† of a normal Hausdorff space X and an open subset A of X such that A is normal and $\operatorname{Ind} A = 1$ but $\operatorname{Ind} X = 0$.

Problem. Does the open subset theorem hold for every completely normal Hausdorff space?

† See 'Local dimension of normal spaces' by the author, to appear.

6. Čech's covering theorem for finite dimensional spaces

Čech ends his paper (**3**) with a proof that a finite-dimensional perfectly normal space admits a certain kind of covering by the closures of disjoint open sets. Theorem 5 below is an extension of Čech's result to totally normal spaces.

[6.1] *Let R be a completely normal space satisfying (b_n) and let* $\operatorname{Ind} R \leqslant n$. *Let U be open, S closed, and U_0 open in S with $U_0 \subset U$ and* $\operatorname{Ind}(\overline{U}_0 - U_0) \leqslant n-1$. *Then there exists V open in R with*

$$U_0 \subset V \subset U, \qquad V \cap S = U_0, \qquad (\overline{V} - V) \cap S = \overline{U}_0 - U_0$$

and $$\operatorname{Ind}(\overline{V} - V) \leqslant n-1.$$

Proof. Let $W = R-(S-U_0)$ and $Y = R-(\overline{U}_0-U_0)$; then W and Y are open and $U_0 \subset W \subset Y$. Since U_0 is closed in the normal space Y and $U_0 \subset W \cap U \subset Y$, there exists H open in Y with

$$U_0 \subset H \subset \overline{H} \cap Y \subset W \cap U.$$

By (b_n), $\operatorname{Ind} Y \leqslant n$, and hence there exists V open in Y and hence open in R with $U_0 \subset V \subset H$ and $\operatorname{Ind}(\overline{V} \cap Y - V) \leqslant n-1$.

Then $V \subset H \subset U$ and $U_0 \subset V \cap S \subset W \cap S = U_0$. Hence $V \cap S = U_0$.

And $$\overline{V} \cap S \cap Y \subset \overline{H} \cap S \cap Y \subset W \cap U \cap S = U_0;$$

hence $$\overline{U}_0 \subset \overline{V} \cap S \subset U_0 \cup (R-Y) = \overline{U}_0.$$

Thus $\overline{V} \cap S = \overline{U}_0$ and

$$(\overline{V} - V) \cap S = \overline{V} \cap S - V \cap S = \overline{U}_0 - U_0.$$

The completely normal space $\overline{V} - V$ is the union of the disjoint sets $\overline{U}_0 - U_0$ and $\overline{V} \cap Y - V$, where $\overline{U}_0 - U_0$ is closed and both sets have $\operatorname{Ind} \leqslant n-1$. Hence, by [2.2], $\operatorname{Ind}(\overline{V} - V) \leqslant n-1$. This completes the proof.

THEOREM 5. *Let R be a totally normal space, or more generally any completely normal space satisfying condition (b_N) for all $N = 0, 1, \ldots$. Let S be closed in R with* $\operatorname{Ind} S \leqslant n$. *Let $U_1, \ldots, U_m$ be open in R with $S \subset \bigcup_{\nu=1}^{m} U_\nu$.*

Then there exist V_i $(1 \leqslant i \leqslant (n+1)m)$ *open in* R *with the properties:*

(i) $\overline{V}_i \subset U_\nu$ *for* $1 \leqslant \nu \leqslant m$, $(n+1)(\nu-1)+1 \leqslant i \leqslant (n+1)\nu$;

(ii) $\bigcup\limits_{i=1}^{(n+1)m} \overline{V}_i \supset S$;

(iii) $V_i \cap V_j = 0$ *for* $1 \leqslant i < j \leqslant (n+1)m$;

(iv) $\operatorname{Ind}((\overline{V}_i - V_i) \cap S) \leqslant n-1$ *for* $1 \leqslant i \leqslant (n+1)m$;

(v) *if* $2 \leqslant r \leqslant n+2$ *and if* $i_1,\ldots, i_r$ *is any combination* (*without repetition*) *of indices* $1, 2,\ldots, (n+1)m$, *then* $\bigcap\limits_{s=1}^{r} \overline{V}_{i_s} \subset S$ *and* $\operatorname{ind} \bigcap\limits_{s=1}^{r} \overline{V}_{i_s} \leqslant n-r+1$.

Proof. Using [3.1], [3.3], and [6.1] above in place of Čech's propositions 23, 19, and 24.1, Čech's proof applies with trivial modifications.

REFERENCES

1. R. H. Bing, 'Metrization of topological spaces', *Canadian J. Math.* 3 (1951) 175–86.
2. L. E. J. Brouwer, 'Über den natürlichen Dimensionsbegriff', *J. f. Math.* 142 (1913) 146–52.
3. E. Čech, 'Dimense dokonale normálních prostorů', *Rozpravy České Akad.* II, 42 (1932), no. 13; 'Sur la dimension des espaces parfaitement normaux', *Bull. int. Acad. Prague* 33 (1932) 38–55.
4. J. Dieudonné, 'Une généralization des espaces compacts', *J. de Math.* 23 (1944) 65–76.
5. O. V. Lokutzievski, 'On the dimension of bicompacta', *Doklady Akad. Nauk SSSR.* 67 (1949) 217–19.

A PROPERTY OF BESSEL FUNCTIONS

By H. G. AP SIMON (*Oxford*)

[Received 2 September 1953]

1. This work originally arose in an attempt to decrease the amount of computation required in a problem in heat conduction involving expressions of the form

$$J_n(ax)Y_n(bx)-Y_n(ax)J_n(bx) \tag{1}$$

in which a and b were known positive constants and x the dependent variable; it was desired to expand (1) as a series in x in the hope that it was sufficiently highly convergent to allow truncation after three or four terms. Although the zeros of (1) have been extensively studied, I can find no explicit statement of the series expansion, and the form it takes may be of some interest.

2. It is convenient to work with modified Bessel functions† and use the notation

$$D_{n,n}(x,y) = (-)^n\{I_n(x)K_n(y)-K_n(x)I_n(y)\}; \tag{2}$$

the result may then be stated as

$$D_{n,n}(2t\cosh k, 2t\sinh k) = \sum_{r=0}^{\infty} \frac{Q_r^n(\cosh 2k)t^{2r}}{r!\,(r+n)!}, \tag{3}$$

where Q_r^n is the associated Legendre function of the second kind, of degree r and order n.

3. From the differential properties of the Bessel functions we find that

$$F(t) \equiv D_{n,n}(2t\cosh k, 2t\sinh k)$$

satisfies the differential equation

$$t^3F^{\text{iv}}+4t^2F'''-(4n^2-1)(tF''-F')-8(\cosh 2k)(t^3F''+2t^2F')+16t^3F = 0. \tag{4}$$

Considering possible series solutions of (4) of the form

$$\sum_{r=0}^{\infty} \frac{a_r t^{2r}}{r!\,(r+n)!},$$

we find that this is a solution of (4) if a_r satisfies

$$(r-n+1)a_{r+1}-(2r+1)(\cosh 2k)a_r+(r+n)a_{r-1} = 0,$$

† The definition taken of $K_n(x)$, the modified Bessel function of the second kind, is that of (**1**), § 3.7, and not that of (**2**), § 17.71.

one solution of which is $Q_r^n(\cosh 2k)$. Hence a solution of (4) is

$$G(t) = \sum_{r=0}^{\infty} \frac{Q_r^n(\cosh 2k)t^{2r}}{r!\,(r+n)!}. \tag{5}$$

4. From the differential equation

$$(z^2-1)\frac{d^2w}{dz^2}+2z\frac{dw}{dz}-\left(n(n+1)+\frac{m^2}{z^2-1}\right)w = 0$$

satisfied by $Q_n^m(z)$, we find that, writing $Q_n^m \equiv Q_n^m(\cosh 2k)$,

$$(m-n)(m+n+1)Q_n^m+2(m+1)(\coth 2k)Q_n^{m+1}+Q_n^{m+2} = 0,$$

and so in particular (from known particular values of Q_n^m)

$$Q_0^0 = \log(\coth k),$$

$$Q_0^m = \tfrac{1}{2}(-)^m(m-1)!\,\{\coth^m k - \tanh^m k\} \quad (m \geqslant 1),$$

$$Q_1^0 = (\cosh 2k)\log(\coth k)-1,$$

$$Q_1^1 = (\sinh 2k)\log(\coth k) - \coth 2k,$$

$$Q_1^m = \tfrac{1}{2}(-)^m(m-2)!\,\{m(\coth^m k + \tanh^m k)- -(\cosh 2k)(\coth^m k - \tanh^m k)\} \quad (m \geqslant 2).$$

5. With the aid of (6) we then have

$$\begin{aligned} G(0) &= Q_0^n(\cosh 2k)/n! &&= F(0),\\ G'(0) &= 0 &&= F'(0),\\ G''(0) &= 2Q_1^n(\cosh 2k)/(n+1)! &&= F''(0),\\ G'''(0) &= 0 &&= F'''(0); \end{aligned}$$

and so the solutions $F(t)$, $G(t)$ of (4) are the same, and the result (3) has been proved.

[This work was sponsored by the Research and Armament Development Division of the Fairey Aviation Company Limited.]

REFERENCES

1. G. N. Watson, *Theory of Bessel Functions* (Cambridge, 2nd ed., 1944).
2. E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge, 4th ed., 1946).

ON THE COVERING OF LATTICE POINTS BY CONVEX REGIONS

By D. B. SAWYER (*Achimota*)

[Received 15 May 1953]

1. Introduction

THE purpose of this paper is to investigate a minimal problem which is suggested by a classical theorem of Minkowski. Minkowski's theorem states that every closed central convex region in the Euclidean plane, with centre at the origin O, and with area at least 4, contains a point of integral coordinates other than O, i.e. a point of the integral lattice in the plane. We consider here a closed central convex region K which is such that, however it is displaced in the plane—a displacement consisting of a translation and a rotation—a point of the integral lattice is covered. Our object is to find the lower bound of the area, $A(K)$, of K. The following result is proved.

THEOREM. $$A(K) \geqslant \tfrac{4}{3},$$

with strict inequality unless K is congruent to the region K^* given by

$$|y| \leqslant \tfrac{3}{4}-x^2, \qquad |x| \leqslant \tfrac{1}{2}.$$

It is easily seen that $A(K^*) = \frac{4}{3}$; also, as will be shown in § 5, K^* is itself a region K, and therefore the number $\frac{4}{3}$ is the true lower bound of $A(K)$.

The problem is transformed and reduced to some extent in § 2, some lemmas are established in § 3, and the theorem is proved in § 4.

2. Reduction of the problem

Notation. We use a vector notation: if X and Y are point sets, then $X+Y$ denotes the set of points $\mathfrak{x}+\mathfrak{y}$, where $\mathfrak{x} \in X$ and $\mathfrak{y} \in Y$.

We denote by $X(x,y,\theta)$ the set obtained by giving X a translation represented by the vector (x,y) and then rotating the set through an angle θ about the point of coordinates (x,y).

We suppose in all that follows, without any loss of generality, that the region K has its centre at O. We denote by d, δ the lengths of the greatest and least diameters of K. The integral lattice will be denoted by Γ, and the general point of Γ by $\mathfrak{g}$. The unit square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ will be denoted by S.

Quart. J. Math. Oxford (2), 4 (1953), 284–92

LEMMA 1. $$A(K) \geqslant \tfrac{1}{2}(d^2-1)^{\frac{1}{2}}+\tfrac{1}{2}\sin^{-1}(1/d). \tag{1}$$

Proof. Parallel tac-lines exist at the extremities of the least diameter of K; if $\delta < 1$, these would be less than unit distance apart, so that K could be displaced into the strip $0 < x < 1$ which contains no point of Γ. Hence $\delta \geqslant 1$, and K contains the circle with centre O and radius $\frac{1}{2}$. But also K contains a line segment of length d with centre at O. Hence, since K is convex, K contains the convex closure of the segment and the circle, a capped circle whose area is given by the right-hand side of (1). This proves the result.

Consider now the region $K(\frac{1}{2},\frac{1}{2},\theta)$. If for some value of θ this contains a point $\mathfrak{g}$ which is not a vertex of S, then

$$(\tfrac{1}{2}d)^2 \geqslant (\tfrac{1}{2})^2+(\tfrac{3}{2})^2, \qquad d \geqslant \sqrt{10},$$

and so, by (1), $$A(K) \geqslant \tfrac{3}{2}+\tfrac{1}{2}\sin^{-1}(1/\sqrt{10}) > \tfrac{4}{3}.$$

Hence we suppose that for each value of θ the region $K(\frac{1}{2},\frac{1}{2},\theta)$ contains no points of Γ other than vertices of S.

By our initial assumption, $K(\frac{1}{2},\frac{1}{2},\theta)$ contains some point $\mathfrak{g}$ for each value of θ. Thus $K(\frac{1}{2},\frac{1}{2},\theta)$ contains a vertex of S and, by symmetry in the point $(\frac{1}{2},\frac{1}{2})$, the opposite vertex of S also. One pair of opposite vertices of S is therefore contained in $K(\frac{1}{2},\frac{1}{2},0)$ and so the other pair is contained in $K(\frac{1}{2},\frac{1}{2},\frac{1}{2}\pi)$. Hence for some value of θ between 0 and $\frac{1}{2}\pi$ all four vertices of S are contained in $K(\frac{1}{2},\frac{1}{2},\theta)$. By the convexity of K it follows that $K(\frac{1}{2},\frac{1}{2},\theta)$ then contains S itself.

Thus K contains a unit square with O as centre. By means of a rotation, if necessary, we may suppose that K contains the square $|x| \leqslant \frac{1}{2}$, $|y| \leqslant \frac{1}{2}$.

LEMMA 2. $$A(K) \geqslant \tfrac{1}{2}\{1+(d^2-1)^{\frac{1}{2}}\}. \tag{2}$$

Proof. Let the coordinates of the extremities of the greatest diameter of K be (x_1,y_1), $(-x_1,-y_1)$. The area of the convex closure of the unit square in K and these extremities is given by

$$|y_1|+\tfrac{1}{2} \quad \text{if } |x_1| \leqslant \tfrac{1}{2},$$

and by $$|x_1|+\tfrac{1}{2} \quad \text{if } |y_1| \leqslant \tfrac{1}{2}.$$

Now $x_1^2+y_1^2 = \frac{1}{4}d^2$, and so, if $|x_1| \leqslant \frac{1}{2}$, we have

$$|y_1| \geqslant \tfrac{1}{2}(d^2-1)^{\frac{1}{2}}.$$

This proves (2) when $|x_1| \leqslant \frac{1}{2}$, and similarly when $|y_1| \leqslant \frac{1}{2}$.

If $|x_1| > \frac{1}{2}$ and $|y_1| > \frac{1}{2}$, the area of the convex closure is

$$|x_1|+|y_1|.$$

Since $|y_1| < \frac{1}{2}(d^2-1)^{\frac{1}{2}}$ when $|x_1| > \frac{1}{2}$, and $|x_1| < \frac{1}{2}(d^2-1)^{\frac{1}{2}}$ when $|y_1| > \frac{1}{2}$, we have

$$(|x_1|-|y_1|)^2 < \{\tfrac{1}{2}(d^2-1)^{\frac{1}{2}}-\tfrac{1}{2}\}^2,$$

and so

$$\begin{aligned}(|x_1|+|y_1|)^2 &= 2(|x_1|^2+|y_1|^2)-(|x_1|-|y_1|)^2\\ &> \tfrac{1}{2}d^2-\tfrac{1}{4}(d^2-1)+\tfrac{1}{2}(d^2-1)^{\frac{1}{2}}-\tfrac{1}{4}\\ &= \tfrac{1}{4}\{(d^2-1)^{\frac{1}{2}}+1\}^2.\end{aligned}$$

This completes the proof.

LEMMA 3. *Let L be any set of points. The set $L+\mathfrak{x}$ contains a point of Γ for every point $\mathfrak{x}$ of the plane if and only if the set $L+\Gamma$ contains the square S.*

Proof. If $L+\Gamma$ contains S, then $L+\Gamma$ contains the whole plane. Hence, if $\mathfrak{x}$ is any point of the plane, $-\mathfrak{x}$ may be represented as $\mathfrak{l}-\mathfrak{g}$ where $\mathfrak{l} \in L$ and $\mathfrak{g} \in \Gamma$. This implies that $\mathfrak{l}+\mathfrak{x}$ is the point $\mathfrak{g}$, so that $L+\mathfrak{x}$ contains $\mathfrak{g}$.

Reversing the argument, if for an arbitrary point $\mathfrak{x}$, $L+\mathfrak{x}$ contains $\mathfrak{g}$, then $-\mathfrak{x}$ may be represented as $\mathfrak{l}-\mathfrak{g}$ and so $-\mathfrak{x}$ is contained in $L-\mathfrak{g}$, and therefore in $L+\Gamma$. Hence S is contained in $L+\Gamma$.

By (2), if $d \geqslant 2$, we have $A(K) \geqslant \frac{1}{2}(1+\sqrt{3}) > \frac{4}{3}$, so that we may suppose $d < 2$. This immediately implies that the intersection $\{K(0,0,\theta)+\mathfrak{g}\} \cap S$ is null unless $\mathfrak{g}$ is a vertex of S. But, by Lemma 3, writing $K(0,0,\theta)$ for L, and recalling that $K(0,0,\theta)+\mathfrak{x}$ always contains a point $\mathfrak{g}$, we see that S is covered by $K(0,0,\theta)+\Gamma$. Hence S is covered by the four regions $K(0,0,\theta)$, $K(1,0,\theta)$, $K(0,1,\theta)$, $K(1,1,\theta)$ alone. It will eventually be shown that this property of K is sufficient to establish the conclusion of the theorem.

3. The main lemmas

Let α, β be constants with $\alpha \geqslant \beta > 0$. We consider the relationship between two variables u, v satisfying

$$\phi(u,v) \equiv u^2+v^2+2\alpha(v-u)-2\beta^2 = 0. \tag{3}$$

Let u take a fixed value with $|u| < \beta$. For large positive or negative v, $\phi(u,v)$ is positive; also

$$\phi(u,-\beta) = u^2-\beta^2-2\alpha(u+\beta) = (u+\beta)(u-\beta-2\alpha) < 0,$$

$$\phi(u,u) = 2(u^2-\beta^2) < 0,$$

while

$$\phi(u,\beta) = u^2-\beta^2-2\alpha(u-\beta) = (u-\beta)(u+\beta-2\alpha) > 0.$$

Thus one root of the equation $\phi(u,v) = 0$, regarded as an equation in v, is less than $-\beta$ and the other lies between u and β.

An immediate consequence of this is that, if we restrict ourselves to values of u and v lying between $-\beta$ and β, the equation (3) defines v as a single-valued function of u, and moreover we have

$$(u-\alpha)\,du+(v+\alpha)\,dv = 0, \tag{4}$$

$$\frac{dv}{du} = \frac{\alpha-u}{\alpha+v} > 0. \tag{5}$$

We also notice that when u takes the extreme values $-\beta$, $+\beta$, the corresponding values of v are also $-\beta$, $+\beta$.

A further consequence of the above work is that, commencing with a value u_0 such that $|u_0| < \beta$, and using the recursion relation

$$\phi(u_n, u_{n+1}) = 0,$$

we can define a sequence $\{u_n\}$ which is strictly increasing, and such that $|u_n| < \beta$ for all n.

LEMMA 4. *If $f(u)$ is a function, continuous in the interval $-\beta \leqslant u \leqslant \beta$, and such that*

$$f(u)+f(v) \geqslant 0 \tag{6}$$

whenever

$$\phi(u,v) = 0, \quad |u| < \beta, \ |v| < \beta, \tag{7}$$

then

$$\int_{-\beta}^{\beta} f(u)\,du \geqslant 0,$$

with strict inequality unless there is always equality at (6).

Proof. Defining v as a function of u by means of (7), we have $f(u)+f(v)$ a continuous function of u on the interval $-\beta \leqslant u \leqslant \beta$. Using (6) and recalling that $\alpha \geqslant \beta$, we have

$$\int_{-\beta}^{\beta} (f(u)+f(v))(\alpha-u)\,du \geqslant 0, \tag{8}$$

and, since the integrand is continuous, there is strict inequality unless there is always equality at (6). But in virtue of (4) and (5) we may rewrite (8) as

$$\int_{-\beta}^{\beta} f(u)(\alpha-u)\,du + \int_{-\beta}^{\beta} f(v)(\alpha+v)\,dv \geqslant 0,$$

i.e.

$$2\alpha \int_{-\beta}^{\beta} f(u)\,du \geqslant 0.$$

LEMMA 5. *If $g(u)$ is a function continuous in the interval $-\beta \leqslant u \leqslant \beta$, and such that, for u, v satisfying* (7),

$$g(u)+g(v) \geqslant v-u+2h, \tag{9}$$

where h is a constant, then

$$\int\limits_{-\beta}^{\beta} g(u)\,du \geqslant 2\beta h+\frac{2\beta^3}{3\alpha}, \tag{10}$$

with strict inequality unless

$$g(u) = g_0(u) \equiv h+(\beta^2-u^2)/2\alpha.$$

Proof. We write

$$f(u) = g(u)-g_0(u) = g(u)-(\beta^2-u^2)/2\alpha-h.$$

Then, by (9),

$$f(u)+f(v) \geqslant v-u+2h-(2\beta^2-u^2-v^2)/2\alpha-2h = 0,$$

so that, by Lemma 4,

$$\int\limits_{-\beta}^{\beta} g(u)\,du = \int\limits_{-\beta}^{\beta} g_0(u)\,du + \int\limits_{-\beta}^{\beta} f(u)\,du$$

$$\geqslant \int\limits_{-\beta}^{\beta} g_0(u)\,du = 2\beta h+\frac{2\beta^3}{3\alpha}.$$

There is strict inequality here unless $f(u)+f(v) = 0$ for all u, v satisfying (7).

Suppose that $f(u)+f(v) = 0$ for all u, v satisfying (7) but that $f(u_0) \neq 0$ for some u_0. Then, forming a sequence $\{u_n\}$ by the recursion

$$\phi(u_n, u_{n+1}) = 0,$$

in the way examined earlier, we have

$$f(u_{2m}) = f(u_0), \qquad f(u_{2m+1}) = -f(u_0)$$

for all positive m. This implies that $f(u)$ is of unbounded variation, which is false, since $f(u)$ is continuous on the closed interval $-\beta \leqslant u \leqslant \beta$.

Hence there is strict inequality in (10) unless $f(u) \equiv 0$, in which case $g(u) = g_0(u)$.

LEMMA 6. *If*

$$0 \leqslant h \leqslant \tfrac{1}{6}, \qquad \alpha = \tfrac{1}{2}+h, \qquad \beta = \{\tfrac{1}{2}-(\tfrac{1}{2}+h)^2\}^{\frac{1}{2}},$$

then (i) $\alpha \geqslant \beta > 0$,

(ii) $h+4\beta h+4\beta^3/3\alpha \geqslant \frac{1}{3}$,

with strict inequality in (ii) *except when $h = 0$.*

Proof.

(i) $\alpha^2-\beta^2 = 2(\tfrac{1}{2}+h)^2-\tfrac{1}{2} \geqslant 0, \qquad \beta^2 \geqslant \tfrac{1}{2}-(\tfrac{1}{2}+\tfrac{1}{6})^2 = \tfrac{1}{18} > 0.$

(ii)
$$\begin{aligned} h+4\beta h+4\beta^3/3\alpha &= h+\beta(12h\alpha+4\beta^2)/3\alpha \\ &= h+\beta(1+2h+8h^2)/3\alpha \geqslant h+\tfrac{2}{3}\beta, \end{aligned}$$

with strict inequality unless $h = 0$. Now

$$4\beta^2 = 1-4h-4h^2 \geqslant 1-6h+9h^2$$

provided that $2h-13h^2 \geqslant 0$, and so, if $0 \leqslant h \leqslant \tfrac{2}{13}$, we have

$$h+\tfrac{2}{3}\beta \geqslant h+\tfrac{1}{3}(1-3h) = \tfrac{1}{3}.$$

If $\tfrac{2}{13} \leqslant h \leqslant \tfrac{1}{6}$, then

$$\frac{1+2h+8h^2}{2\alpha} = 1+\frac{8h^2}{1+2h} \geqslant 1+\frac{32}{221} > \frac{8}{7},$$

since $h^2/(1+2h)$ is an increasing function. We thus have

$$h+4\beta h+4\beta^3/3\alpha > h+\tfrac{16}{21}\beta.$$

But

$$\begin{aligned} (16\beta)^2-49(1-3h)^2 &= 15+38h-697h^2 \\ &\geqslant 15+\tfrac{76}{13}-\tfrac{697}{36} > 0. \end{aligned}$$

Hence, if $\tfrac{2}{13} \leqslant h \leqslant \tfrac{1}{6}$, then

$$h+\tfrac{16}{21}\beta > h+\tfrac{1}{3}(1-3h) = \tfrac{1}{3}.$$

This proves the result.

4. Proof of the theorem

The unit square $|x| \leqslant \tfrac{1}{2}$, $|y| \leqslant \tfrac{1}{2}$ is contained in K. Hence, if we write

$$\sup_{\mathbf{x}\in K} x = \tfrac{1}{2}+h, \qquad \sup_{\mathbf{x}\in K} y = \tfrac{1}{2}+h',$$

we have $h \geqslant 0$, $h' \geqslant 0$, and, by remarking that the convex closure of the unit square and the extreme points is contained in K, we have

$$A(K) \geqslant 1+h+h'.$$

There is no loss of generality in supposing $h \leqslant h'$ and then

$$A(K) \geqslant 1+2h,$$

so that we need be concerned only with the range $0 \leqslant h \leqslant \tfrac{1}{6}$.

Since K lies in the half-plane $x \leqslant \tfrac{1}{2}+h$, $K(0,0,\theta)$ lies in the half-plane

$$x\cos\theta+y\sin\theta \leqslant \tfrac{1}{2}+h.$$

This half-plane does not contain the point $(\tfrac{1}{2},\tfrac{1}{2})$, the centre of S, when

$$\cos\theta+\sin\theta > 1+2h, \tag{11}$$

and by symmetry, in this case, $K(0,0,\theta)$ and $K(1,1,\theta)$ together cannot cover S; in fact the strip

$$\tfrac{1}{2}+h < x\cos\theta+y\sin\theta < \cos\theta+\sin\theta-\tfrac{1}{2}-h \tag{12}$$

lies between $K(0,0,\theta)$ and $K(1,1,\theta)$.

When (11) holds, the line

$$x\cos\theta+y\sin\theta = \tfrac{1}{2}+h \tag{13}$$

separates O and the centre of S. It lies at a distance $\frac{1}{2}+h \geqslant \frac{1}{2}$ from O and so meets the boundary of S at points within a distance $\frac{1}{2}$ from the vertices $(1,0)$ and $(0,1)$: that is, at points belonging to $K(1,0,\theta)$ and $K(0,1,\theta)$ respectively. Also, when (11) holds, the line (13) contains no interior points of $K(0,0,\theta)$ and $K(1,1,\theta)$, and so all points common to this line and S are in either $K(1,0,\theta)$ or $K(0,1,\theta)$, since otherwise there would be points of the strip (12) not covered. It therefore follows, since the intersections of $K(1,0,\theta)$ and $K(0,1,\theta)$ with the line (13) are closed segments, that *some* point of the line (13) belongs to both $K(1,0,\theta)$ and $K(0,1,\theta)$.

Now K is defined by relations of the form

$$y \leqslant \tfrac{1}{2}+g(x), \qquad -y \leqslant \tfrac{1}{2}+g(-x), \qquad |x| \leqslant \tfrac{1}{2}+h,$$

where $g(x)$ is continuous in the interval $-\frac{1}{2}-h \leqslant x \leqslant \frac{1}{2}+h$, and $g(x) \geqslant 0$ when $|x| \leqslant \frac{1}{2}$. Hence all points of $K(1,0,\theta)$ satisfy

$$y\cos\theta-(x-1)\sin\theta \leqslant \tfrac{1}{2}+g\{(x-1)\cos\theta+y\sin\theta\},$$

and all points of $K(0,1,\theta)$ satisfy

$$(1-y)\cos\theta+x\sin\theta \leqslant \tfrac{1}{2}+g\{-x\cos\theta-(y-1)\sin\theta\}.$$

Since some point satisfying (13) also satisfies both of these inequalities, we have on addition

$$\cos\theta+\sin\theta \leqslant 1+g(h+\tfrac{1}{2}-\cos\theta)+g(\sin\theta-h-\tfrac{1}{2}),$$

i.e.
$$g(u)+g(v) \geqslant v-u+2h,$$

where
$$u = h+\tfrac{1}{2}-\cos\theta, \qquad v = \sin\theta-h-\tfrac{1}{2}. \tag{14}$$

From (14), (11) we have

$$u^2+v^2+2\alpha(v-u)-2\beta^2 = 0 \tag{15}$$

and
$$v-u > 0, \tag{16}$$

where
$$\alpha = \tfrac{1}{2}+h, \qquad \beta^2 = \tfrac{1}{2}-(\tfrac{1}{2}+h)^2.$$

From (15), (16),
$$u^2+v^2 < 2\beta^2,$$

so that either $|u| < \beta$ or $|v| < \beta$. But the initial examination of the relationship $\phi(u,v) = 0$ shows that, if $|u| < \beta$ and (16) holds, then

$|v| < \beta$ also. Similarly, since (15), (16) remain unaltered if we replace u, v by $-v, -u$ respectively, if $|v| < \beta$ then also $|u| < \beta$. Hence both

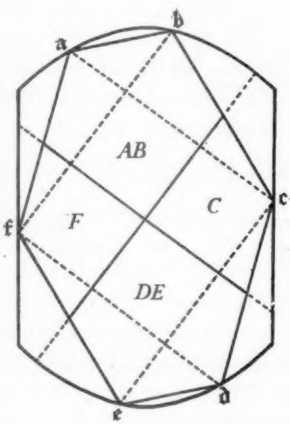
$$|u| < \beta, \qquad |v| < \beta. \tag{17}$$

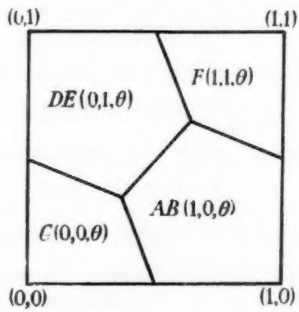
Further, any values of u and v which satisfy (15) and (17) satisfy

$$(u-h-\tfrac{1}{2})^2+(v+h+\tfrac{1}{2})^2 = 1,$$

$$u-h-\tfrac{1}{2} < 0, \qquad v+h+\tfrac{1}{2} > 0,$$

and so are of the form (14).





Thus the conditions of Lemma 5 are satisfied, and, using Lemmas 5 and 6, and remarking that $\beta \leqslant \frac{1}{2}$, we have

$$A(K) \geqslant 1+h+2\int\limits_{-\frac{1}{2}}^{\frac{1}{2}} g(x)\,dx \geqslant 1+h+2\int\limits_{-\beta}^{\beta} g(x)\,dx \geqslant \tfrac{4}{3}.$$

There is strict inequality unless $h = 0$ and $g(x) = g_0(x)$, i.e. unless $K = K^*$.

5. The region K^*

It remains to be shown that K^* is itself a K. By Lemma 3, it is enough to show that the regions $K^*(0,0,\theta)$, $K^*(1,0,\theta)$, $K^*(0,1,\theta)$, $K^*(1,1,\theta)$ cover S for each value of θ. The figure formed by the regions for some value of θ is the same as the figure formed for the value $\theta+\frac{1}{2}\pi$ except for a rotation of $\frac{1}{2}\pi$ of the whole configuration about the centre of S. Because of this, we need prove only that S is covered for $0 \leqslant \theta \leqslant \frac{1}{2}\pi$.

We consider the hexagon H, inscribed in K^*, with vertices

$$\mathfrak{a}\,(\tfrac{1}{2}-\cos\theta, \cos\theta-\tfrac{1}{2}\cos 2\theta), \qquad \mathfrak{d}\,(-\tfrac{1}{2}+\cos\theta, -\cos\theta+\tfrac{1}{2}\cos 2\theta),$$
$$\mathfrak{b}\,(\sin\theta-\tfrac{1}{2}, \tfrac{1}{2}\cos 2\theta+\sin\theta), \qquad \mathfrak{e}\,(-\sin\theta+\tfrac{1}{2}, -\tfrac{1}{2}\cos 2\theta-\sin\theta),$$
$$\mathfrak{c}\,(\tfrac{1}{2}, \cos\theta-\sin\theta-\tfrac{1}{2}\cos 2\theta), \qquad \mathfrak{f}\,(-\tfrac{1}{2}, -\cos\theta+\sin\theta+\tfrac{1}{2}\cos 2\theta).$$

This is symmetrical in O. The chords $\mathfrak{ac}$, $\mathfrak{fd}$, $\mathfrak{bf}$, $\mathfrak{ce}$ are each of unit length, and $\mathfrak{ac}$, $\mathfrak{fd}$ are parallel to the line

$$x\sin\theta+y\cos\theta = 0, \tag{18}$$

while $\mathfrak{bf}$, $\mathfrak{ce}$ are parallel to the line

$$x\cos\theta-y\sin\theta = 0. \tag{19}$$

The lines (18), (19) divide H into four subsets, which we may denote AB, C, DE, F according to the vertices of H contained in them. Then (see figure above) it is immediate that S is the union

$$C(0,0,\theta) \cup AB(1,0,\theta) \cup DE(0,1,\theta) \cup F(1,1,\theta),$$

so that *a fortiori* S is contained in

$$K^*(0,0,\theta) \cup K^*(1,0,\theta) \cup K^*(0,1,\theta) \cup K^*(1,1,\theta).$$

This completes the investigation.

I am very grateful to the referee for a number of comments, and particularly for the present proof of Lemma 4, which is much shorter than the original one.

IRREDUCIBLE CONVEX BODIES

By KATHLEEN OLLERENSHAW (*Manchester*)

[Received 1 June 1953]

1. Introduction

A STAR body S is said to be *irreducible* if, for every star body T contained in but different from S, $\Delta(T) < \Delta(S)$, where $\Delta(T)$, $\Delta(S)$ are the *critical*† *determinants* of T, S. In this note I prove the theorems

THEOREM 1. *The generalized cylinder*

$$f(x_1, x_2) \leqslant 1, \quad |x_3| \leqslant 1, \ldots, |x_n| \leqslant 1,$$

where $f(x_1, x_2) \leqslant 1$ *defines any irreducible convex star domain, is irreducible.*

THEOREM 2. *The n-dimensional sphere for all* $n \leqslant 5$ *is irreducible.*

Mahler [(**2**) 336, Theorem C] and Rogers [(**4**) Theorem 4] have shown that a *necessary* condition for S to be irreducible is that every point on the boundary belongs to a critical lattice. The generalized cylinder on a convex irreducible base (which, of course, includes the three-dimensional cylinder on a convex irreducible star domain as base and the n-dimensional parallelepiped) and the n-dimensional sphere are the only known convex bodies of more than two dimensions that satisfy this necessary condition; of these the critical determinants of the n-spheres ($n > 8$) have not yet been found.

The only‡ star body of more than two dimensions previously known to be irreducible is the 3-sphere.

† The critical determinant $\Delta(S)$ of a star body S is the lower bound of the determinants of the lattices Λ *admissible* for S, i.e. the lattices Λ with no point other than the origin in the interior of S.

‡ Mahler (**2**) (341, Theorem F) proved that a 3-dimensional *convex* star body S is irreducible if every point of its boundary belongs to a critical lattice of S with just twelve points on the boundary. The 3-sphere has this property and is therefore irreducible. Mahler also stated that the circular cylinder and 3-cube were irreducible on the grounds that every point on their boundaries belonged to a critical lattice with just twelve points on the boundaries. This is incorrect: critical lattices containing a point on a generator of a convex cylinder other than on the base itself all contain at least fourteen points on the boundary of the cylinder; and all critical lattices of the cube contain at least fourteen points on the boundary.

2. I use Rogers' concept of *irreducible points* (**4**). A point P of a star body S is called *reducible* if there is a star body T contained in S but not containing P for which $\Delta(T) = \Delta(S)$. A point of S which is not reducible is called *irreducible*. Rogers proves that

a star body S (of the finite type) is irreducible if and only if every point on its boundary is irreducible;

if S is a star body of the finite type, the set Σ of irreducible points of S is closed;

a point P on the boundary of a star body S (of the finite type) is irreducible if and only if, for every $\epsilon > 0$, there exists a lattice Λ^* with $d(\Lambda^*) < \Delta(S)$ such that the only points of Λ^* in the interior of S are the points O, $\pm P^*$, where $|P-P^*| < \epsilon$.

Rogers' results agree with Mahler's earlier work in terms of *free* critical lattices. Mahler calls a critical lattice Λ *free* if, for every point P of Λ on the boundary of S, there exists a neighbouring lattice Λ^* satisfying the conditions above.

3. Proof of Theorem 1

Denote by K the irreducible convex star domain $f(x_1, x_2) \leqslant 1$ and by C the generalized cylinder defined by $f(x_1, x_2) \leqslant 1$, $|x_3| \leqslant 1,\ldots, |x_n| \leqslant 1$, where there is no loss of generality in supposing the axes rectangular. Then† $\Delta(C) = \Delta(K)$. We have to show that every point on the boundary of C is irreducible. Mahler [(**3**) 695, Theorem 3] has shown that the boundary of K is either a parallelogram or a strictly convex region with a continuously turning tangent. Since the set of irreducible points of C is closed, it is sufficient to prove irreducible the points for which either (i) $|x_r| < 1$ $(r = 3,\ldots, n)$ and $f(x_1, x_2) = 1$, where, if K is a parallelogram, (x_1, x_2) is not the mid-point of a side of K or a vertex of K; or

$$\text{(ii)} \quad |x_{r \neq r'}| < 1, \ |x_{r'}| = 1$$

where r' is one of the $r = 3,\ldots, n$, and $f(x_1, x_2) < 1$, where x_1, x_2 are not both zero.

† The result $\Delta(C) = \Delta(K)$ is true for a generalized cylinder on any convex base. When the base is a circle, the result is due to Mahler [*Quart. J. of Math.* (Oxford), 17 (1946), 16–18: footnote on p. 16]. The general result follows by use of Mahler's method or by a result of Hlawka 'Ausfüllung und Überdeckung durch Zylinder', *Anz. Oster. Akad. Wiss. Wien, Math.-Nat. Kl.* 85 (1948), 116–9; or see *Math. Rev.* 11 (1950), 12] combined with a result of L. Fejes Tôth [*Acta Sci. Math.* (Szeged) 12 (1950), 62–7; or see Rogers, *Acta Math.* 86 (1951), 309–21]. This information was kindly given me by Dr. Rogers. It was at his suggestion that I proved Theorem 1 in the present form: originally I proved the result only for the three-dimensional cylinder and the n-cube.

By symmetry and by a suitable choice of the (x_1, x_2)-axes, these points are typified by (i) the point $P = (x_1', x_2', \alpha_3,..., \alpha_n)$, say, where

$$0 \leqslant \alpha_3,..., \alpha_n < 1$$

and $P' = (x_1', x_2')$ is any point of the boundary of K (other than the mid-point of a side or a vertex of K if K is a parallelogram); (ii) the point

$$Q = (\alpha, 0, \alpha_3, \alpha_4,..., \alpha_{n-1}, 1),$$

where $0 \leqslant \alpha_3,..., \alpha_{n-1} < 1$, and $Q' = (\alpha, 0)$ $(\alpha > 0)$ is an inner point of K other than the centre. In §§ 4–8 I prove irreducible the points P so defined. In §§ 9–11 I deal with the points Q.

4. Consider first the point $P' = (x_1', x_2')$ on the boundary of K and such that, if K is a parallelogram, P' is not the mid-point of a side or a vertex. Mahler† has shown that, corresponding to every such point P', there exists a point P'' on the boundary of K such that $\pm(P''-P')$ is also on the boundary of K. Moreover P', P'' so defined generate a critical lattice of K with just these six points $\pm P'$, $\pm P''$, $\pm(P''-P')$ on the boundary of K.

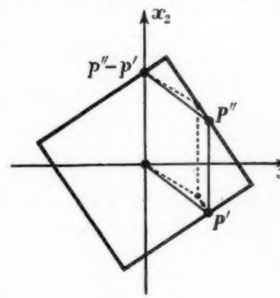


FIG. 1.

Let the coordinates of P'' on the boundary of K be (x_1'', x_2'') and choose the (x_1, x_2)-axes so that $x_1' = x_1'' = \bar{x}_1 > 0$ (say), $x_2''-x_2' = \bar{x}_2 > 0$ (say); i.e. so that $P' = (\bar{x}_1, x_2')$, $P'' = (\bar{x}_1, x_2'+\bar{x}_2)$. By hypothesis,

$$(P''-P') = (0, \bar{x}_2)$$

is on the boundary of K, and, by Mahler's result, P', P'' generate a two-dimensional critical lattice L_1 (say) of K, such that

$$d(L_1) = \Delta(K) = \bar{x}_1 \bar{x}_2 > 0.$$

Moreover $\pm P'$, $\pm P''$, $\pm(P''-P')$ are the only points of L_1 on the boundary of K.

Consider the points $P_\epsilon' = (\bar{x}_1-\epsilon, x_2'+\delta)$, $P_\epsilon'' = (\bar{x}_1-\epsilon, x_2'+\bar{x}_2+\delta)$ in the plane of K, where $\epsilon > 0$ is arbitrarily small and $\delta > 0$ is such that P_ϵ'' is on the boundary of K. This choice is clearly possible when K is strictly convex.

From the well-known properties of the critical lattices of a parallelogram it is also possible when K is a parallelogram since P' is not the mid-point of a side or a vertex (Fig. 1). Moreover $\delta \to 0$ as $\epsilon \to 0$ and

$$|P'-P_\epsilon'| = \sqrt{(\epsilon^2+\delta^2)} \to 0 \quad \text{as } \epsilon \to 0.$$

† (3) 101–2, and the properties of the critical lattices of a parallelogram.

If L_1^* is the two-dimensional lattice generated by P'_ϵ, P''_ϵ in the plane of K, then L_1^* is in the neighbourhood of L_1; 0, $\pm P'_\epsilon$ are inner points of K; $\pm P''_\epsilon$, $\pm(P''_\epsilon - P'_\epsilon) = \pm(P''-P')$ lie on the boundary of K, and, since, by Mahler's result, all points of L_1 other than O, $\pm P'$, $\pm P''$, $\pm(P''-P')$ lie outside K, all points of L_1^*, other than O, $\pm P'_\epsilon$, lie on the boundary of K or outside K.

5. Consider now the lattice Λ_1 (say) containing $P = (\bar{x}_1, x'_2, \alpha_3,..., \alpha_n)$, where $0 \leqslant \alpha_3,..., \alpha_n < 1$, and defined by the equations

$$\left.\begin{aligned}
x_1 &= \bar{x}_1\xi_1 + \bar{x}_1\xi_2 \\
x_2 &= x'_2\xi_1 + (x'_2+\bar{x}_2)\xi_2 \\
x_3 &= \alpha_3\xi_1 + \alpha_3\xi_2 + \xi_3 \\
x_4 &= \alpha_4\xi_1 + \alpha_4\xi_2 + \alpha_4\xi_3 + \xi_4 \\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\
x_n &= \alpha_n\xi_1 + \alpha_n\xi_2 + \alpha_n\xi_3 + \alpha_n\xi_4 + ... + \xi_n
\end{aligned}\right\} (\Lambda_1).$$

This lattice has the determinant $d(\Lambda_1) = \bar{x}_1\bar{x}_2 = \Delta(K)$. It is clearly admissible for C since, by Mahler's result and the definitions of P', P'', if $(\xi_1, \xi_2) \neq (0, 0)$, then $f(x_1, x_2) \geqslant 1$; and, if $(\xi_1, \xi_2) = (0, 0)$, then

$$\max|x_r| \geqslant 1 \quad (r = 3,..., n)$$

except when all ξ_r are zero, i.e. except for the lattice point at the origin. It follows that Λ_1 is a critical lattice of C.

Denote by Λ_1^* the neighbouring lattice defined by the equations

$$\left.\begin{aligned}
x_1 &= (\bar{x}_1-\epsilon)\xi_1 + (\bar{x}_1-\epsilon)\xi_2 - \epsilon\xi_3 - \epsilon\xi_4 - ... - \epsilon\xi_n \\
x_2 &= (x'_2+\delta)\xi_1 + (x'_2+\bar{x}_2+\delta)\xi_2 + \delta\xi_3 + \delta\xi_4 + ... + \delta\xi_n \\
x_3 &= \alpha_3\xi_1 + \alpha_3\xi_2 + \xi_3 \\
x_4 &= \alpha_4\xi_1 + \alpha_4\xi_2 + \alpha_4\xi_3 + \xi_4 \\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\
x_n &= \alpha_n\xi_1 + \alpha_n\xi_2 + \alpha_n\xi_3 + \alpha_n\xi_4 + ... + \xi_n
\end{aligned}\right\} (\Lambda_1^*),$$

where $\epsilon > 0$ is arbitrarily small and δ is defined as in the previous paragraph. Then Λ_1^* contains the point $P_\epsilon = (\bar{x}_1-\epsilon, x'_2+\delta, \alpha_3,..., \alpha_n)$ in the interior of C and such that $|P-P_\epsilon| \to 0$ as $\epsilon \to 0$. By definition the points $\pm(x_1-\epsilon, x'_2+\bar{x}_2+\delta, \alpha_3,..., \alpha_n)$, $\pm(0, \bar{x}_2, 0,..., 0)$ are on the boundary of C. The determinant of Λ_1^* is

$$d(\Lambda_1^*) = \bar{x}_2[\bar{x}_1 - \epsilon(1-\alpha_3)(1-\alpha_4)...(1-\alpha_n)] < \bar{x}_1\bar{x}_2 = \Delta(K) = \Delta(C).$$

We have thus to show that P, $\pm P_\epsilon$ are the only points of Λ_1^* in the interior of C.

6. Since Λ_1, Λ_1^* are neighbouring lattices, the only points of Λ_1^* other than O, $\pm P_\epsilon$ which could be in the interior of C are those in the neighbourhood of points of Λ_1 which lie on the boundary of C. Moreover, since x_r ($r = 3,\dots, n$) are identical in the equations defining Λ_1, Λ_1^*, and, since Λ_1 is admissible for C, all lattice points of Λ_1, Λ_1^* for which $\max|x_r| \geqslant 1$ ($r \geqslant 3$) lie on the boundary of C or outside C. We have thus to consider only those points of Λ_1^* in the neighbourhood of points of Λ_1 for which $\max|x_r| < 1$, $f(x_1, x_2) = 1$ and we have already seen that $f(x_1, x_2) = 1$ only if $(\xi_1, \xi_2) = \pm(1, 0)$, $\pm(0, 1)$, or $\pm(1, -1)$.

In the equations for Λ_1, if $(\xi_1, \xi_2) = \pm(1, -1)$, then $\max|x_r| < 1$ for $r \geqslant 3$ only if all $x_r = 0$, i.e. if the lattice point is one of $\pm(0, \bar{x}_2, 0,\dots, 0)$ which by definition is on the boundary of C.

7. Consider then the equations for Λ_1 and suppose that

$$(\xi_1, \xi_2) = \pm(1, 0) \quad \text{or} \quad \pm(0, 1), \quad \max|x_r| < 1.$$

We see that $|x_3| < 1$ only if either (i) $\xi_3 = 0$ or (ii) $\xi_1+\xi_2+\xi_3 = 0$. If (i) $\xi_3 = 0$, then $|x_4| < 1$ only if $\xi_4 = 0$ or $\xi_1+\xi_2+\xi_4 = 0$; if (ii) $\xi_1+\xi_2+\xi_3 = 0$, then $|x_4| < 1$ only if $\xi_4 = 0$. We thus have $|x_3| < 1$, $|x_4| < 1$ only if either (*a*) $\xi_3 = \xi_4 = 0$ or (*b*) $\xi_1+\xi_2+\xi_3+\xi_4 = 0$. If (*a*) $\xi_3 = \xi_4 = 0$, then $|x_5| < 1$ only if $\xi_5 = 0$ or $\xi_1+\xi_2+\xi_5 = 0$; if (*b*) $\xi_1+\xi_2+\xi_3+\xi_4 = 0$ then $|x_5| < 1$ only if $\xi_5 = 0$. We thus have $|x_3| < 1$, $|x_4| < 1$, $|x_5| < 1$ only if either $\xi_3 = \xi_4 = \xi_5 = 0$ or

$$\xi_1+\xi_2+\xi_3+\xi_4+\xi_5 = 0.$$

Continuing in this way we find eventually that, if $(\xi_1, \xi_2) = \pm(1, 0)$ or $\pm(0, 1)$, then $\max|x_r| < 1$ ($r \geqslant 3$) only if either all

$$\xi_r = 0 \ (r = 3,\dots, n) \quad \text{or} \quad \sum \xi \equiv \xi_1+\xi_2+\xi_3+\xi_4+\dots+\xi_n = 0.$$

We now consider points of Λ_1^* for which these conditions on the ξ hold. If $\sum \xi = 0$, the equations of Λ_1, Λ_1^* become identical. Hence points of Λ_1^* other than the origin for which $\sum \xi = 0$ are not in the interior of C. The points of Λ_1^* given by $(\xi_1, \xi_2,\dots, \xi_n) = \pm(0, 1, 0,\dots, 0)$ are by hypothesis on the boundary of C. The points of Λ_1^* given by $\pm(1, 0, 0,\dots, 0)$ are of course $\pm P_\epsilon$. We have thus shown that the points O, $\pm P_\epsilon$ are the only points of Λ_1^* in the interior of C, as was to be proved.

8. By symmetry and permutation of the coordinates x_r, we have thus shown that every point on a generator of C is irreducible. If K is a parallelogram, by a suitable permutation of the coordinates x, every point on the boundary of C (now an n-dimensional parallelepiped) can be thought of as a point on a generator. Theorem 1 is thus proved for the particular case when K is a parallelogram. It remains therefore to

prove irreducible the points Q defined in § 3, when K is a strictly convex star domain with a continuous tangent.

9. Before discussing the irreducibility of the point

$$Q = (\alpha, 0, \alpha_3, \alpha_4, ..., \alpha_{n-1}, 1),$$

where $0 \leqslant \alpha_3, ..., \alpha_{n-1} < 1$, and $Q' = (\alpha, 0)$ $(\alpha > 0)$ is an inner point of K, it is necessary to prove certain properties of the plane region K.

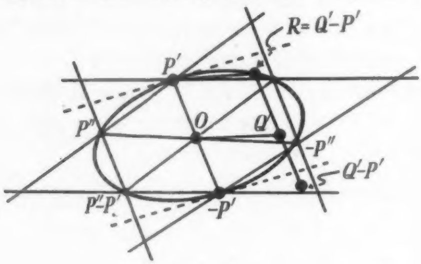


FIG. 2.

Let $P'R$ be that chord of K (Fig. 2) which is equal and parallel in the same sense to OQ' and has positive x_2-coordinate. Since K is strictly convex, $P'R$ is uniquely defined. Moreover, since R lies 'below' the tangent at P', $Q'-P'$ lies 'below' the tangent at $-P'$, and, since K has a continuous tangent, the point $Q'-P'$ lies outside K. Hence, no point $Q' \pm hP'$ other than Q', where h is positive integer or zero, is in the interior of K.

Let P' on the boundary of K have coordinates (x_1', x_2') $(x_2' > 0)$ and let $P'' = (x_1'', x_2'')$, say, be that point of the boundary of K such that $(P''-P')$ is also a point of the boundary and $(x_1' x_2'' - x_2' x_1'') > 0$. Then P'' is uniquely defined and $\Delta(K) = x_1' x_2'' - x_2' x_1''$. The points P'', Q' lie on opposite sides of OP'. Since K is strictly convex and Q' is an inner point of K, R is an inner point of the triangle with vertices at P', $P'-P''$, $2P'-P''$. Hence $Q' = R-P'$ is an inner point of the triangle

$$O(-P'')(P'-P'').$$

Hence the area $OP'(P'-P'') = \frac{1}{2}\Delta(K)$ is greater than the area

$$OP'Q' = \tfrac{1}{2}\alpha x_2',$$

i.e. $\Delta(K) > \alpha x_2' > 0$. Moreover all points $Q' \pm hP' - kP''$, where h is a positive integer or zero and k $(\neq 0)$ is a positive integer, lie outside the six-pointed star $[S]$, say, formed by producing the sides of the hexagon with vertices $\pm P'$, $\pm P''$, $\pm(P''-P')$. Since K is strictly convex, $[S]$ contains K, and it follows that no point $Q' \pm hP' - kP''$ other than Q', where h, k are positive integers or zero, is in the interior of K.

10. We can now return to n dimensions. Let us consider the lattice Λ_2 (say) defined by the equations

$$\left.\begin{array}{l} x_1 = x_1'\xi_1 + x_1''\xi_2 + \alpha\xi_3 + \alpha\xi_4 + \ldots + \alpha\xi_n \\ x_2 = x_2'\xi_1 + x_2''\xi_2 \\ x_3 = \xi_3 + \alpha_3\xi_4 + \ldots + \alpha_3\xi_n \\ x_4 = \xi_4 + \ldots + \alpha_4\xi_n \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_n = \xi_n \end{array}\right\} (\Lambda_2),$$

This lattice contains the point Q and has the determinant

$$d(\Lambda_2) = (x_1'x_2'' - x_1''x_2') = \Delta(K).$$

It is clearly admissible, since $\max|x_r| < 1$ $(r \geqslant 3)$ only if all $\xi_r = 0$ $(r = 3,\ldots, n)$ and, by the definition of P', P'', no point other than the origin of the two-dimensional lattice in the plane of K generated by P', P'' is in the interior of C. The lattice Λ_2 is thus a critical lattice of C.

Let Λ_2^* be the neighbouring lattice defined by the equations

$$\left.\begin{array}{l} x_1 = x_1'\xi_1 + x_1''\xi_2 + \alpha\xi_3 + \alpha\xi_4 + \ldots + \alpha\xi_n \\ x_2 = x_2'\xi_1 + x_2''\xi_2 \\ x_3 = \xi_3 + \alpha_3\xi_4 + \ldots + \alpha_3\xi_n \\ x_4 = \xi_4 + \ldots + \alpha_4\xi_n \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_n = \epsilon\xi_2 - \epsilon\xi_3 - \epsilon\xi_4 - \ldots + (1-\epsilon)\xi_n \end{array}\right\} (\Lambda_2^*),$$

where $\epsilon > 0$ is arbitrarily small. Then Λ_2^* contains the point

$$Q_\epsilon = (\alpha, 0, \alpha_3, \alpha_4,\ldots, \alpha_{n-1}, 1-\epsilon), \text{ say,}$$

in the interior of C and $|Q - Q_\epsilon| = \epsilon$. By definition the points

$$\pm(x_1'', x_2'', 0,\ldots, 0), \quad \pm(x_1'' - x_1', x_2'' - x_2', 0,\ldots, 0)$$

of Λ_2, Λ_2^* are on the boundary of C. The determinant of Λ_2^* is

$$d(\Lambda_2^*) = \Delta(K) - \epsilon[\Delta(K) - \alpha x_2'](1-\alpha_3)(1-\alpha_4)\ldots(1-\alpha_{n-1}) < \Delta(K) = \Delta(C).$$

We have thus to show that O, $\pm Q_\epsilon$ are the only points of Λ_2^* in the interior of C.

11. We proceed as in § 6, considering only points of Λ_2^* in the neighbourhood of points of Λ_2 which lie on the boundary of C. Since the equations for x_s $(s = 1,\ldots, n-1)$ are identical in the definitions of Λ_2, Λ_2^*, and since Λ_2 is admissible for C, all lattice points of Λ_2, Λ_2^* for which

$$\max[f(x_1, x_2), |x_t|] \geqslant 1 \quad (2 < t < n)$$

lie on or outside C. We have thus to consider only those points of Λ_2^* in the neighbourhood of points of Λ_2 for which

$$\max[f(x_1, x_2), |x_t|] < 1 \quad (2 < t < n), \qquad |x_n| = 1.$$

In the equations for Λ_2, $|x_n| = 1$ only if $\xi_n = \pm 1$. Then $|x_{n-1}| < 1$ only if either (i) $\xi_{n-1} = 0$ or (ii) $\xi_{n-1}+\xi_n = 0$. If (i) $\xi_{n-1} = 0$, then $|x_{n-2}| < 1$ only if either $\xi_{n-2} = 0$ or $\xi_{n-2}+\xi_n = 0$; if (ii) $\xi_{n-1}+\xi_n = 0$, then $|x_{n-2}| < 1$ only if $\xi_{n-2} = 0$. We thus have $|x_{n-1}| < 1$, $|x_{n-2}| < 1$ only if either (*a*) $\xi_{n-1} = \xi_{n-2} = 0$ or (*b*) $\xi_n+\xi_{n-1}+\xi_{n-2} = 0$. Continuing in this way as in § 7 we find eventually that, if $|x_n| = 1$ in the equations of Λ_2, then $\max|x_t| < 1$ $(2 < t < n)$ only if either all $\xi_t = 0$ $(2 < t < n)$ or $\xi_3+...+\xi_{n-1}+\xi_n = 0$.

We now consider points of Λ_2^* for which these conditions on the ξ hold. If $\xi_3+...+\xi_{n-1}+\xi_n = 0$, then $x_1 = x_1'\xi_1+x_1''\xi_2$, $x_2 = x_2'\xi_1+x_2''\xi_2$ in the equations of both Λ_2, Λ_2^*. But, by hypothesis, $f(x_1, x_2) < 1$, so that $\xi_1 = \xi_2 = 0$ and the point of Λ_2^* has $x_n = 1$ and lies on the boundary of C or outside C. The equations for the points $[Q]$, say, of Λ_2^* which satisfy the conditions $\xi_n = \pm 1$, $\xi_t = 0$, are

$$x_1 = x_1'\xi_1+x_1''\xi_2+\alpha\xi_n, \quad x_2 = x_2'\xi_1+x_2''\xi_2, \quad x_3 = \alpha_3\xi_n, \quad x_4 = \alpha_4\xi_n,$$
$$..., \quad x_{n-1} = \alpha_{n-1}\xi_n, \quad x_n = \epsilon\xi_2+(1-\epsilon)\xi_n.$$

By the symmetry of C we need consider only the points of $[Q]$ for which $\xi_n = 1$. If $\xi_2 > 0$, then $x_n = 1$ and the points lie outside C. It remains to consider just those points of $[Q]$ for which $\xi_2 \leqslant 0$. But the points

$$x_1 = x_1'\xi_1+x_1''\xi_2+\alpha, \qquad x_2 = x_2'\xi_1+x_2''\xi_2 \quad (\xi_2 \leqslant 0)$$

in the plane of K are just those points $Q'\pm hP'-kP''$ (h, k positive integers or zero) discussed in § 9. There it was shown that no such point other than Q' is in the interior of K. It follows that $f(x_1, x_2) \geqslant 1$ for all points of the set $[Q]$ other than the points given by $(\xi_1, \xi_2) = (0, 0)$, i.e. the points

$$\pm Q_\epsilon = \pm(\alpha, 0, \alpha_3, \alpha_4, ..., \alpha_{n-1}, 1).$$

We have thus shown that O, $\pm Q_\epsilon$ are the only points of Λ_2^* in the interior of C. It follows that Q and all the points typified by Q are irreducible. This together with the results of § 8 shows that every point of the boundary of C is irreducible and so C itself is irreducible.

This completes the proof of Theorem 1.

12. Proof of Theorem 2

Let S_n be the n-sphere defined by $\sum x_r^2 \leqslant 1$ $(r = 1, 2, ..., n)$ and choose the axes so that P, any point of the surface of S_n, has coordinates

$x_1 = 1$, $x_s = 0$ $(s = 2,..., n)$. Write $\epsilon > 0$, where ϵ is arbitrarily small, and let P^* be the point $x_1 = 1-\epsilon$, $x_s = 0$ in the neighbourhood of P and in the interior of S_n.

Consider the equations

$$x_1 = (1-\epsilon)\xi_1+\tfrac{1}{2}(1-\epsilon)\xi_2+\tfrac{1}{2}(1-\epsilon)\xi_3,$$

$$x_2 = \tfrac{1}{2}(3+2\epsilon-\epsilon^2)^{\frac{1}{2}}\xi_2+\frac{(1+2\epsilon-\epsilon^2)\xi_3}{2(3+2\epsilon-\epsilon^2)^{\frac{1}{2}}}+\frac{\xi_4}{(3+2\epsilon-\epsilon^2)^{\frac{1}{2}}},$$

$$x_3 = \frac{(2+2\epsilon-\epsilon^2)^{\frac{1}{2}}}{(3+2\epsilon-\epsilon^2)^{\frac{1}{2}}}\xi_3+\frac{\xi_4}{2(2+2\epsilon-\epsilon^2)^{\frac{1}{2}}(3+2\epsilon-\epsilon^2)^{\frac{1}{2}}}+\frac{(3+2\epsilon-\epsilon^2)^{\frac{1}{2}}\xi_5}{2(2+2\epsilon-\epsilon^2)^{\frac{1}{2}}},$$

$$x_4 = \frac{(1+2\epsilon-\epsilon^2)^{\frac{1}{2}}\xi_4}{(2+2\epsilon-\epsilon^2)^{\frac{1}{2}}}+\frac{(1+2\epsilon-\epsilon^2)^{\frac{1}{2}}\xi_5}{2(2+2\epsilon-\epsilon^2)^{\frac{1}{2}}},$$

$$x_5 = \frac{1}{\sqrt{2}}\xi_5.$$

Define the n-dimensional lattice Λ_n^* $(n \leqslant 5)$ by these equations with $x_r = 0 = \xi_r$ $(r > n)$. Then we have:†

$$d(\Lambda_2^*) = \tfrac{1}{2}(1-\epsilon)(3+2\epsilon-\epsilon^2)^{\frac{1}{2}} = \tfrac{1}{2}(3-4\epsilon-2\epsilon^2...)^{\frac{1}{2}} < \tfrac{1}{2}\sqrt{3} = \Delta(S_2),$$
$$d(\Lambda_3^*) = \tfrac{1}{2}(1-\epsilon)(2+2\epsilon-\epsilon^2)^{\frac{1}{2}} = \tfrac{1}{2}(2-2\epsilon-3\epsilon^2...)^{\frac{1}{2}} < \tfrac{1}{2}\sqrt{2} = \Delta(S_3),$$
$$d(\Lambda_4^*) = \tfrac{1}{2}(1-\epsilon)(1+2\epsilon-\epsilon^2)^{\frac{1}{2}} = \tfrac{1}{2}(1-4\epsilon^2...)^{\frac{1}{2}} < \tfrac{1}{2} = \Delta(S_4),$$
$$d(\Lambda_5^*) = \tfrac{1}{4}\sqrt{2}(1-\epsilon)(1+2\epsilon-\epsilon^2)^{\frac{1}{2}} = \tfrac{1}{4}\sqrt{2}(1-4\epsilon^2...)^{\frac{1}{2}} < \tfrac{1}{4}\sqrt{2} = \Delta(S_5).$$

Moreover we find that

$$\begin{aligned}\sum_{r=(1\text{ to }5)} x_r^2 &= (1-\epsilon)^2\xi_1(\xi_1+\xi_2+\xi_3)+\xi_2(\xi_2+\xi_3+\xi_4)+\\ &\qquad +\xi_3(\xi_3+\xi_4+\xi_5)+\xi_4(\xi_4+\xi_5)+\xi_5^2\\ &= (1-\epsilon)^2X+Y, \text{ say,}\end{aligned}$$

where X, Y are integer forms and Y is positive-definite in $\xi_2,..., \xi_5$. Let X_n, Y_n be the forms obtained from X, Y by putting $\xi_r = 0$ for $r > n$. Then, if we put $\xi_r = 0$ for $r > n$, we obtain

$$\sum_{1,...,n} x_r^2 = (1-\epsilon)^2X_n+Y_n.$$

Let Λ be the lattice Λ^* with $\epsilon = 0$. Then $d(\Lambda) = \Delta(S_n)$, and $\sum x_r^2$ is a positive-definite integer form, and Λ is admissible for S_n. Hence Λ is a critical lattice of S_n and the only points of Λ $(\epsilon = 0)$ on the surface are those given by (integer) values of $(\xi_1, \xi_2,..., \xi_n)$ for which

$$\sum x_r^2 = X_n+Y_n = 1.$$

† The values of $\Delta(S)$ $(n \leqslant 8)$ are now well known. A full account of work done in this field, together with an extensive bibliography, is given by Coxeter (**1**).

Hence the only points of the neighbouring lattice Λ^* ($\epsilon > 0$) which could be in the interior of S_n other than O are those for which $X_n+Y_n = 1$, i.e. since Y_n is positive definite, $X_n = 1$ or $X_n \leqslant 0$. But

$$\sum x_r^2 = (1-\epsilon)^2X_n+Y_n = X_n+Y_n-\epsilon(2-\epsilon)X_n.$$

If $X_n+Y_n = 1$ and $X_n = 1$, i.e. if

$$\xi_1 = \pm 1, \qquad \xi_2 = \xi_3 = \ldots = \xi_n = 0,$$

then $\sum x_r^2 < 1$, i.e. the point $\pm P^*$ is in the interior of S. If $X_n+Y_n = 1$ and $X_n \leqslant 0$, then $\sum x_r^2 \geqslant 1$.

It follows that $d(\Lambda_n^*) < \Delta(S_n)$ ($n \leqslant 5$), and the only points of Λ_n^* in the interior of S_n are 0, $\pm P^*$, where $|P-P^*| = \epsilon$. Hence P, and so every point of the surface of S_n ($n \leqslant 5$), is irreducible.

This completes the proof of Theorem 2.

REFERENCES

1. H. S. M. Coxeter, *Canadian J. of Math.* 3 (1951), 391–441.
2. K. Mahler, *Proc. Kon. Ned. Akad. v. Wetensch.* (Amsterdam) 49 (1946), 331–43.
3. —— ibid. 50 (1947), 98–118; 692–703.
4. C. A. Rogers, ibid. 379–83.

ON THE REPRESENTATION OF REAL NUMBERS BY PRODUCTS OF RATIONAL NUMBERS

A. OPPENHEIM (*Singapore*)

[Received 17 July 1953]

1. CANTOR (**1**) showed that any real number $x > 1$ is uniquely expressible in the form

$$x = \left(1+\frac{1}{a_1}\right)\left(1+\frac{1}{a_2}\right)\left(1+\frac{1}{a_3}\right)..., \tag{1.1}$$

where the a_i are integers such that

$$a_1 \geqslant 1,\ a_2 \geqslant a_1^2,\ a_3 \geqslant a_2^2,\ \dots. \tag{1.2}$$

To obtain a_i he employed the algorithm given by

$$1+\frac{1}{a_1} < x = x_1 \leqslant 1+\frac{1}{a_1-1}, \qquad a_1 = \left[\frac{x_1}{x_1-1}\right],$$

$$x_2 = \frac{a_1 x_1}{a_1+1} > 1, \qquad a_2 = \left[\frac{x_2}{x_2-1}\right], \qquad \text{etc.} \tag{1.3}$$

Conversely, given integers a_i which satisfy (1.2) and such that at least one $a_i \geqslant 2$, the product in (1.1) converges to a number $x > 1$ and the a_i are the integers associated with x by the algorithm.

Cantor also proved that x is rational if and only if, from some point on, each a_i is the square of its immediate predecessor, i.e. if

$$a_{i+1} = a_1^2 \quad (i \geqslant i_0).$$

From such products as (1.1) remarkable approximations were obtained.

It is interesting to note a generalization of (1.1) into

$$x = \prod_1^\infty \left(1+\frac{n_i}{d_i}\right), \tag{1.4}$$

where $n_i \geqslant 1$, $d_i \geqslant 1$ are integers. As it stands, (1.4) is too general. It is essential to restrict the n_i and d_i in some way.

In this note I consider the following case:

$$n_i = c_1 c_2 ... c_i, \quad c_i \text{ a positive integer.} \tag{1.5}$$

An algorithm will be employed to determine the d_i. It is remarkable that criteria analogous to those of Cantor for $c_i \equiv 1$ can be given in the general case (1.5) to cover uniqueness of the expansion and to settle

Quart. J. Math. Oxford (2), 4 (1953), 303-7.

rationality or irrationality. As an example, if $n_i \equiv c \geqslant 1$, there is an expansion in which

$$d_{i+1} \geqslant (d_i-1)(d_i+c)+1, \qquad d_1 \geqslant 1 \quad (i = 1, 2, \ldots).$$

The number x is rational if and only if equality holds for all i from some point on. The complete statements will be given later.

2. We define the integer $d_1 \geqslant 1$ by the inequalities

$$1+\frac{n_1}{d_1} < x = x_1 \leqslant 1+\frac{n_1}{d_1-1},$$

so that

$$d_1 = 1+\left[\frac{n_1}{x_1-1}\right].$$

Write

$$x_1 = \left(1+\frac{n_1}{d_1}\right)x_2,$$

so that $x_2 > 1$. Define the integer $d_2 \geqslant 1$ by

$$1+\frac{n_2}{d_2} < x_2 \leqslant 1+\frac{n_2}{d_2-1},$$

so that

$$d_2 = 1+\left[\frac{n_2}{x_2-1}\right].$$

In this way we define a sequence of integers $d_i \geqslant 1$. Now clearly

$$\frac{n_2}{d_2} < x_2-1 = \frac{(x_1-1)d_1-n_1}{d_1+n_1} \leqslant \frac{n_1}{(d_1-1)(d_1+n_1)},$$

so that

$$d_2 > \frac{n_2}{n_1}(d_1-1)(d_1+n_1).$$

Thus we have defined integers $d_i \geqslant 1$ and numbers $x_i > 1$ such that

$$x_1 = x, \qquad d_i = 1+\left[\frac{n_i}{x_i-1}\right] \quad (i = 1, 2, \ldots),$$

$$x_{i+1} = \frac{d_i x_i}{d_i+n_i}, \qquad d_{i+1} > \frac{n_{i+1}}{n_i}(d_i-1)(d_i+n_i). \tag{2.1}$$

Plainly, if $d_i = 1$ for some i, then $d_j = 1$ for $j = 1, \ldots, i$. But

$$x > \left(1+\frac{n_1}{d_1}\right)\left(1+\frac{n_2}{d_2}\right)\cdots\left(1+\frac{n_i}{d_i}\right),$$

$$x > (1+n_1)(1+n_2)\ldots(1+n_i) \geqslant 2^i,$$

so that i is *bounded*. Hence at most the first i_0 of the d_i, where

$$i_0 = [\log x/\log 2],$$

can be unity and the rest of the d_i must exceed unity.

Consider now

$$d_{i+1}-(d_i-1) > (d_i-1)\left(\frac{n_{i+1}}{n_i}d_i+n_{i+1}-1\right) > 0$$

if $d_i \geqslant 2$. Hence $d_{i+1} > d_i-1$, $d_{i+1} \geqslant d_i$.

But, if $d_{i+1} = d_i \geqslant 2$, then necessarily

$$0 < (d_i-1)\left(\frac{n_{i+1}}{n_i}d_i+n_{i+1}-1\right) < 1.$$

Hence $$n_{i+1} = 1, \qquad n_i > d_i(d_i-1) > 1.$$

Changing i into $i+1$ we see therefore that

$$d_{i+2} > d_{i+1}.$$

Note that, if $d_{i+1} = d_i \geqslant 2$, then also

$$1+\frac{n_i}{d_i} > d_i, \qquad 1+\frac{n_i}{d_i} \geqslant 1+d_i \geqslant 3.$$

Thus the number of times that $d_{i+1} = d_i \geqslant 2$ can occur is bounded (being at most $[\log x/\log 3]$). Hence d_i is a non-decreasing sequence of positive integers up to $i = i_0$ and thereafter d_i is steadily increasing to infinity.

The series of positive terms

$$\sum u_i, \quad u_i = n_i/d_i,$$

is such therefore that

$$\frac{u_{i+1}}{u_i} = \frac{n_{i+1}d_i}{d_{i+1}n_i} < \frac{d_i}{(d_i-1)(d_i+n_i)},$$

which tends to 0 as $i \to \infty$. Hence $\sum u_i$ and therefore also $\prod (1+u_i)$ is convergent, its value being x.

To sum up: we have proved

THEOREM 1. *Given integers $n_i \geqslant 1$ $(i = 1, 2,...)$, there exist for any $x > 1$ integers $d_i \geqslant 1$ $(i = 1, 2,...)$ such that*

$$x = \prod_{i=1}^{\infty}\left(1+\frac{n_i}{d_i}\right) \tag{2.2}$$

and $$1 \leqslant d_1 \leqslant d_2 \leqslant d_3 \ldots \to \infty,$$

$$d_{i+1} > \frac{n_{i+1}}{n_i}(d_i-1)(d_i+n_i), \tag{2.3}$$

the integers d_i being chosen by (2.1).

3. The question now arises: given a convergent product (2.2) in which the integers d_i satisfy (2.3), is it the case that the expansion is that associated with x and numerators n_i by means of (2.1)?

I consider here the case

$$n_i = c_1 c_2 \ldots c_i, \qquad c_i \text{ a positive integer,} \tag{3.1}$$

and I prove

THEOREM 2. *If* (3.1) *holds and if the integers $d_i \geqslant 1$ satisfy*

$$d_{i+1} \geqslant c_{i+1}(d_i - 1)(d_i + n_i) + 1 \quad (i = 1, 2, \ldots), \tag{3.2}$$

and at least one $d_i \geqslant 2$, then the product

$$\prod_1^\infty (1+u_i) = x < \infty \tag{3.3}$$

is such that the d_i are associated with x and numerators n_i by (2.1).

It will be enough to prove that

$$1+\frac{n_i}{d_i} < x_i = \prod_{j=1}^{\infty}\left(1+\frac{n_j}{d_j}\right) \leqslant 1+\frac{n_i}{d_i - 1}. \tag{3.4}$$

The left-hand inequality is trivial. That on the right requires proof only when $d_i \geqslant 2$. Such a d_i exists and then, by (3.2), $d_j \geqslant 2$ for all $j \geqslant i$.

Now, for $d_i \geqslant 2$,

$$1+\frac{n_i}{d_i-1} = \left(1+\frac{n_i}{d_i}\right)\left(1+\frac{n_i}{(d_i-1)(d_i+n_i)}\right) \geqslant \left(1+\frac{n_i}{d_i}\right)\left(1+\frac{n_{i+1}}{d_{i+1}-1}\right), \tag{3.5}$$

with equality if and only if

$$d_{i+1} - 1 = c_{i+1}(d_i - 1)(d_i + n_i).$$

Repeated application of (3.5) yields

$$\prod_{j=i}^{i+m}\left(1+\frac{n_j}{d_j}\right) < 1+\frac{n_i}{d_i-1},$$

and (3.4) follows on letting $m \to \infty$. Thus Theorem 2 is proved.

4. Rationality of x given by (3.3)

THEOREM 3. *If equality holds in* (3.2) *for all $i \geqslant i_0$, then x is rational.*

For the argument used in (3.5) shows that

$$x_i = 1+\frac{n_i}{d_i - 1} \quad (d_i \geqslant 2)$$

for each $i \geqslant i_0$: thus x_i is rational, so that x, being a rational multiple of x_i, is also rational.

More interesting and naturally not so obvious is the converse given in

THEOREM 4. *If x is rational, then, from some point on,*

$$d_{i+1} = c_i(d_i - 1)(d_i + c_1 c_2 \ldots c_i) + 1; \tag{4.1}$$

and x, given by (3.3), (3.2), *is rational if and only if* (4.1) *holds from some point on.*

Suppose that

$$x_i = p_i/q_i, \quad p_i > q_i > 1, \quad (p_i, q_i) = 1.$$

Then
$$\frac{p_{i+1}}{q_{i+1}} = x_{i+1} = \frac{d_i x_i}{n_i+d_i} = \frac{d_i p_i}{q_i(n_i+d_i)},$$

so that
$$\lambda_i p_{i+1} = d_i p_i, \qquad \lambda_i q_i = (d_i+n_i)q_i, \tag{4.2}$$

where
$$\lambda_i = \{d_i p_i, (d_i+n_i)q_i\} \geqslant 1.$$

Now
$$\frac{p_i}{q_i} = x_i \leqslant 1+\frac{n_i}{d_i-1},$$

so that
$$p_i d_i-(d_i+n_i)q_i \leqslant p_i-q_i. \tag{4.3}$$

From (4.2), (4.3) we obtain

$$1 \leqslant p_{i+1}-q_{i+1} \leqslant \lambda_i(p_{i+1}-q_{i+1}) \leqslant p_i-q_i. \tag{4.4}$$

The sequence of positive integers $\gamma_i = p_i-q_i$ is therefore non-increasing. Hence for $i = i_0$ it reaches a least value $\gamma \geqslant 1$; for all $i \geqslant i_0$ it follows from (4.4) that

$$\gamma_{i+1} = \gamma_i = \gamma, \qquad \lambda_i = 1,$$

and that equality holds in (4.3); hence

$$\frac{n_{i+1}}{d_{i+1}-1} = \frac{\gamma}{q_{i+1}} = \frac{\gamma}{(d_i+n_i)q_i} = \frac{n_i}{(d_i-1)(d_i+n_i)},$$

$$d_{i+1}-1 = c_{i+1}(d_i-1)(d_i+n_i) \quad (i_1 \geqslant i_0),$$

which proves (4.1).

The second part of Theorem 4 follows by the uniqueness of the expansion.

5. An example of this algorithm using $n_i \equiv 2$ is worth giving:

$$\sqrt{2} = (1+\tfrac{2}{5})(1+\tfrac{2}{197})(1+\tfrac{2}{7761797})...,$$

which is a particular case of the remarkable expansion, due, as I have since found, to Escott (**2**),

$$\left(\frac{a+3}{a-1}\right)^{\frac{1}{2}} = \left(1+\frac{2}{a_1}\right)\left(1+\frac{2}{a_2}\right)\left(1+\frac{2}{a_3}\right)...$$

where
$$a_1 = a > 1, \qquad a_{i+1} = a_i^3+3a_i^2-3.$$

REFERENCES

1. G. Cantor, 'Zwei Sätze über eine gewisse Zerlegung der Zahlen in unendliche Produkte', *Collected Papers* (1932), 43–50.

2. E. B. Escott, 'Rapid method for extracting square roots', *American Math. Monthly*, 44 (1947), 644–6.

APPROXIMATION TO π BY TRIGONOMETRICAL SURDS

By L. S. GODDARD (*Aberdeen*)

[Received 15 June 1953]

1. VARIOUS approximations to π by means of algebraic numbers have been given, and for some of these geometrical constructions have been described. The simplest is $\sqrt{2}+\sqrt{3} = 3{\cdot}1462...$, and Heiseler (1) has given two constructions which lead (2) to the values

$$\sqrt{(\tfrac{40}{3}-2\sqrt{3})} = 3{\cdot}14153..., \qquad \tfrac{1}{3}(\sqrt{141}-\sqrt{6}) = 3{\cdot}14161....$$

Ramanujan (3), with the aid of certain modular equations, has derived some remarkable approximations, the best being

$$\pi \sim \frac{63}{25}\left(\frac{17+15\sqrt{5}}{7+15\sqrt{5}}\right), \qquad \frac{1}{2\pi\sqrt{2}} \sim \frac{1103}{99^2},$$

which give π correct to 9 and 8 decimal places respectively. Ramanujan's analysis is rather heavy, but he obtained empirically the value

$$\pi \sim \left(9^2+\frac{19^2}{22}\right)^{\frac{1}{4}},$$

which gives π correct to 8 decimals, and gave a geometrical construction for this number.

In the present paper we consider the problem of finding further algebraic approximations for π; some of those found are much more accurate than those previously obtained. Two methods are described for generating sequences of approximations. These are elementary and depend on the principle that, if $f(\sin\theta, \cos\theta)$ is a rational function of $\sin\theta$ and $\cos\theta$ which, when expanded as a power series in θ, is of the form

$$f(\sin\theta, \cos\theta) = \theta + O(\theta^k),$$

where $k > 2$, then by putting $\theta = n\pi/60$, where n is a small integer we obtain

$$\pi \sim \frac{60}{n} f\left(\sin\frac{n\pi}{60}, \cos\frac{n\pi}{60}\right),$$

and, since the sine and cosine of $n\pi/60$ are algebraic numbers (4), it follows that this approximation to π is algebraic. The closeness of the approximation is practically determined by k and by the coefficient of θ^k in the above expansion.

Quart. J. Math. Oxford (2), 4 (1953), 308–13.

2. In the first method we consider the sum

$$f_n(\theta) = \sum_{r=1}^{n} a_r \sin r\theta,$$

where the a_r are chosen so that, in the expansion of $f_n(\theta)$ as a series in θ, the terms involving θ^3, θ^5,..., θ^{2n-1} are zero, and the coefficient of θ is unity. The equations determining the a_r are

$$a_1+2a_2+3a_3+...+na_n = 1,$$

$$a_1+2^{2p+1}a_2+3^{2p+1}a_3+...+n^{2p+1}a_n = 0 \quad (p = 1, 2,..., n-1).$$

Let $\Delta(\alpha_2,..., \alpha_n)$ denote the determinant whose ith row is $(1, \alpha_2^{2i}, \alpha_3^{2i},..., \alpha_n^{2i})$ for $i = 0, 1, 2,..., n-1$. It is easy to show that

$$\Delta(\alpha) = \Delta(\alpha_2,..., \alpha_n) = (-)^{\frac{1}{2}(n-1)(n-2)} \prod_{j=2}^{n} (\alpha_j^2-1) - \prod_{j<k} (\alpha_j^2-\alpha_k^2),$$

and we have $\qquad ra_r = (-)^{r-1}(n!)^2\Delta_r(\alpha)/\Delta(\alpha),$

where $\Delta_r(\alpha)$ is the determinant obtained from $\Delta(\alpha)$ by deleting the last row and the rth column, and where we put $\alpha_s = s$ $(s = 2,..., n)$. After some reduction we find that

$$a_r = \frac{(-)^{r-1}2(n!)^2}{(n-r)!\,(n+r)!\,r} \quad (r = 1, 2,..., n).$$

For the remainder in the expansion we have, by Maclaurin's theorem,

$$R_{2n+1}(\theta) = \frac{\theta^{2n+1}}{(2n+1)!} f_n^{(2n+1)}(\lambda\theta) \quad (0 < \lambda < 1)$$

$$= \pm \frac{\theta^{2n+1}}{(2n+1)!} \sum_{r=1}^{n} r^{2n+1}a_r \cos \lambda r\theta.$$

Hence

$$|R_{2n+1}(\theta)| = \frac{\theta^{2n+1}}{(2n+1)!} \left| \sum_{r=1}^{n} r^{2n+1}a_r \cos \lambda r\theta \right|$$

$$< \frac{\theta^{2n+1}}{(2n+1)!} \sum_{r=1}^{n} r^{2n+1}a_r.$$

Now

$$\sum_{r=1}^{n} r^{2n+1}a_r = \frac{\Delta'}{n!\,\Delta(\alpha)},$$

where Δ' is the determinant whose ith row is $(1, 2^{2i+3},..., n^{2i+3})$ for $i = 0, 1,..., n-1$. Since

$$\Delta' = (n!)^3\Delta(\alpha),$$

it follows that

$$|R_{2n+1}(\theta)| < \frac{(n!)^2\theta^{2n+1}}{(2n+1)!}.$$

By taking $n = 2, 3, 4, \ldots$ we obtain the set of inequalities:†

$$n = 2: \quad |\theta - (\tfrac{4}{3}\sin\theta - \tfrac{1}{6}\sin 2\theta)| < \frac{\theta^5}{30},$$

$$n = 3: \quad |\theta - (\tfrac{3}{2}\sin\theta - \tfrac{3}{10}\sin 2\theta + \tfrac{1}{30}\sin 3\theta)| < \frac{\theta^7}{140},$$

$$n = 4: \quad |\theta - (\tfrac{8}{5}\sin\theta - \tfrac{2}{5}\sin 2\theta + \tfrac{8}{105}\sin 3\theta - \tfrac{1}{140}\sin 4\theta)| < \frac{\theta^9}{630},$$

$$n = 5: \quad |\theta - (\tfrac{5}{3}\sin\theta - \tfrac{10}{21}\sin 2\theta + \tfrac{5}{42}\sin 3\theta - \tfrac{5}{252}\sin 4\theta + \tfrac{1}{630}\sin 5\theta)| < \frac{\theta^{11}}{2772},$$

$$n = 6: \quad |\theta - (\tfrac{12}{7}\sin\theta - \tfrac{15}{28}\sin 2\theta + \tfrac{10}{63}\sin 3\theta - \tfrac{1}{28}\sin 4\theta + \tfrac{2}{385}\sin 5\theta - \tfrac{1}{2772}\sin 6\theta)| < \frac{\theta^{13}}{12012}.$$

. . . .

If, in any of these, we put $\theta = k\pi/60$ where $k = 1, 2, 3, \ldots$ we obtain a sequence of algebraic approximations to π, and an upper bound for the error.

Thus, with $n = 3$, $k = 5$, we obtain

$$\pi \sim \tfrac{1}{10}[9(5\sqrt{6}-2)-43\sqrt{2}] = 3{\cdot}1415867\ldots,$$

which gives π correct to 5 decimal places.

The best result derived from this set of approximations is, putting $n = 6$, $k = 1$,

$$\pi \sim \frac{30}{7}\left[\frac{1}{18}(27s_1+5s_3) - \frac{1}{16}(15s_2+s_4) + \frac{1}{11}\left(\frac{s_5}{5} - \frac{s_6}{72}\right)\right],$$

where

$$\begin{aligned} s_1 &= (\sqrt{6}+\sqrt{2})(\sqrt{5}-1) - 2(\sqrt{3}-1)\sqrt{(5+\sqrt{5})}, \\ s_2 &= \sqrt{(30-6\sqrt{5})} - \sqrt{5} - 1, \\ s_3 &= \sqrt{10}+\sqrt{2} - 2\sqrt{(5-\sqrt{5})}, \\ s_4 &= \sqrt{(10+2\sqrt{5})} - \sqrt{15} + \sqrt{3}, \\ s_5 &= \sqrt{6} - \sqrt{2}, \\ s_6 &= \sqrt{5} - 1. \end{aligned}$$

This expression gives π correct to 18 places of decimals.‡

3. In the second method we use a continued-fraction development. By using a standard transformation for the quotient of two hypergeometric series, the well-known series,

$$\theta = \sin\theta\cos\theta\left(1 + \tfrac{2}{3}\sin^2\theta + \frac{2\,.\,4}{3\,.\,5}\sin^4\theta + \ldots\right),$$

† Since $\tfrac{1}{2}a_r \to (-1)^{r-1}/r$ as $n \to \infty$, the limiting form of $f_n(\theta)$ is the Fourier series

$$\theta = 2(\sin\theta - \tfrac{1}{2}\sin 2\theta + \tfrac{1}{3}\sin 3\theta - \ldots) \quad (0 \leqslant \theta < \pi).$$

‡ The expression may be evaluated numerically by using the values given (to 24 decimal places) by Gray (**5**).

can be transformed [(4), 375] into the continued fraction

$$\theta = \frac{a_1}{1-}\,\frac{a_2}{1-}\,\frac{a_3}{1-}\cdots,$$

where $$a_1 = \sin\theta\cos\theta, \qquad a_{2r} = \frac{2r(2r-1)}{(4r-3)(4r-1)}\sin^2\theta,$$

$$a_{2r+1} = \frac{2r(2r-1)}{(4r-1)(4r+1)}\sin^2\theta \quad (r = 1, 2,\ldots).$$

If p_n/q_n is the nth convergent, we have

$$p_n = p_{n-1} - a_n p_{n-2}, \qquad q_n = q_{n-1} - a_n q_{n-2} \quad (n = 2, 3,\ldots), \tag{1}$$

where $p_0 = 0$, $p_1 = a_1$, $q_0 = 1$, $q_1 = 1$.

Regarding the difference $\theta - p_n/q_n$ we have

LEMMA 1. $$\theta - \frac{p_n}{q_n} < \frac{a_1 a_2 \ldots a_{n+1}}{q_n q_{n+2}}.$$

If we formally expand p_n/q_n into a power series in θ, we have

LEMMA 2.

$$\theta - \frac{p_{2r}}{q_{2r}} = \mu\theta^{4r+1}P_1(\theta), \qquad \theta - \frac{p_{2r+1}}{q_{2r+1}} = \nu\theta^{4r+3}P_2(\theta),$$

where

$$\mu = \frac{2^{4r-2}}{(4r+1)\binom{4r-1}{2r}^2}, \qquad \nu = \frac{(r+1)2^{4r+1}}{(2r+1)(4r+3)\binom{4r+1}{2r}^2},$$

and $P_1(\theta)$, $P_2(\theta)$ are power series with constant term equal to unity.

The proof of Lemma 1 is as follows. The relations (1) lead to

$$p_n q_{n-1} - p_{n-1} q_n = a_1 a_2 \ldots a_n,$$

and it is easily shown that

$$\theta - \frac{p_n}{q_n} = \frac{p_{n+1}q_n - p_n q_{n+1}}{q_n(q_{n+1} - q_n r_{n+1})},$$

where $$r_n = \frac{a_{n+1}}{1-}\,\frac{a_{n+2}}{1-}\cdots.$$

But $r_n > a_{n+1}$, so that

$$q_{n+1} - q_n r_{n+1} > q_{n+1} - q_n a_{n+2} = q_{n+2}.$$

Thus $$\theta - \frac{p_n}{q_n} = \frac{a_1 a_2 \ldots a_{n+1}}{q_n(q_{n+1} - q_n r_{n+1})} < \frac{a_1 a_2 \ldots a_{n+1}}{q_n q_{n+2}}.$$

We use this lemma to deduce Lemma 2. It is observed that q_n is a polynomial in $\sin^2\theta$ with constant term equal to unity; also $a_1 a_2 \ldots a_{n+1}$ involves θ in the form $\sin^{2n+1}\theta\cos\theta$. Thus

$$\theta - p_n/q_n = O(\theta^{2n+1}),$$

and the coefficients μ and ν follow from the values of the a_r.

To obtain approximations for π we now put $\theta = k\pi/60$, as before, and take $k\pi/60 \sim p_n/q_n$. An estimate of the error involved is given by Lemma 2. Since the values of μ and ν are small, even for $r = 1, 2, 3,\ldots$, the error is small.

The values of p_n/q_n for $n = 1, 2, 3,\ldots$ are as follows:

$$p_1/q_1 = \tfrac{1}{2}\sin 2\theta, \qquad p_2/q_2 = \frac{3\sin 2\theta}{2(2+\cos 2\theta)}, \qquad p_3/q_3 = \frac{28\sin 2\theta + \sin 4\theta}{12(3+2\cos 2\theta)},$$

$$p_4/q_4 = \frac{5}{12}\left(\frac{32\sin 2\theta + 5\sin 4\theta}{18+16\cos 2\theta + \cos 4\theta}\right),$$

$$p_5/q_5 = \frac{425\sin 2\theta + 101\sin 4\theta + \sin 6\theta}{60(10+10\cos 2\theta + \cos 4\theta)},$$

$$p_6/q_6 = \frac{7}{20}\left(\frac{375\sin 2\theta + 132\sin 4\theta + 7\sin 6\theta}{200+225\cos 2\theta + 36\cos 4\theta + \cos 6\theta}\right),$$

$$\cdots\cdots\cdots$$

These convergents give rise to a large number of approximations to π. I record only a few. When $\theta = \pi/12$ and $n = 3, 4, 5$, we obtain

$$\pi \sim \frac{28+\sqrt{3}}{2(3+\sqrt{3})} = 3{\cdot}1415608\ldots,$$

$$\pi \sim 5\left(\frac{32+5\sqrt{3}}{37+16\sqrt{3}}\right) = 3{\cdot}14159222\ldots,$$

$$\pi \sim \frac{1}{5}\left(\frac{427+101\sqrt{3}}{21+10\sqrt{3}}\right) = 3{\cdot}141592642\ldots,$$

which give π correct to 4, 5, and 7 places of decimals respectively.

When $\theta = \pi/24$, $n = 5$ we obtain

$$\pi \sim \frac{42}{5}\left(\frac{264+375\sqrt{6}-361\sqrt{2}}{800+225\sqrt{6}+72\sqrt{3}+227\sqrt{2}}\right),$$

which gives π correct to 14 decimal places.

Finally, when $\theta = \pi/60$, $n = 6$, we have

$$\frac{\pi}{21} \sim \frac{375\{\sqrt{(30-6\sqrt{5})}-\sqrt{5}-1\}+132\{\sqrt{(10+2\sqrt{5})}-\sqrt{15}+\sqrt{3}\}+14(\sqrt{5}-1)}{1600+225\{\sqrt{(10-2\sqrt{5})}+\sqrt{15}+\sqrt{3}\}+36\{\sqrt{(30+6\sqrt{5})}+\sqrt{5}-1\}+2\sqrt{(10+2\sqrt{5})}}$$

and this gives π correct to 18 decimals.

REFERENCES

1. A. Heiseler, 'Die Würfelverdoppelung, Winkeldreiteilung und Rectifikation des Kreises durch elementare Näherungskonstruktionen', *Acta Acad. Aboensis, Math. et Phys.* 8 (1935), 8 pp.
2. N. Pipping, 'Drei Geometrische Minaturen', ibid. 18 (1952), 7 pp.
3. S. Ramanujan, 'Modular equations and approximations to π', *Quart. J. of Math.* 45 (1914), 350–72.
4. E. W. Hobson, *Treatise on Plane Trigonometry* (Cambridge, 1921), 74.
5. P. Gray, 'Values of the trigonometrical quadratic surds', *Mess. of Math.* 6 (1877), 105–6.

A CYLINDRICAL CURVE WITH MAXIMUM LENGTH AND MAXIMUM HEIGHT

By M. K. FORT, Jr. (*Athens, Ga.*)

[Received 6 July 1953]

1. Introduction

In a recent paper (**1**), J. W. Green considers curves C which have diameter 2 and which encircle a right circular cylinder of unit radius. Each such curve C turns out to be rectifiable and hence has a length $L(C)$. $H(C)$ is defined to be the minimum distance between two planes perpendicular to the cylinder and containing C between them. Green proves that

$$2\pi \leqslant L(C) \leqslant 2\sqrt{2}\,\pi,$$

$$0 \leqslant H(C) \leqslant \sqrt{2}$$

and shows that these bounds are the best possible. He also proves that there exists a curve C for which $H(C) = \sqrt{2}$. The following two questions are asked but not answered in Green's paper. First, *does there exist a curve C for which $L(C) = 2\sqrt{2}\,\pi$*; and second, *what is the least upper bound for $H(C)+L(C)$?* In this paper I construct a curve C for which $L(C) = 2\sqrt{2}\,\pi$ and $H(C) = \sqrt{2}$. My example thus answers both of the above questions, and shows that it is indeed possible for C to have simultaneously both maximum length and maximum height.

2. Method of construction

It follows from Green's paper that in order to obtain a curve C having the desired properties, it is sufficient to construct a continuous function f on the interval $[0, \frac{1}{2}\pi]$ with the following properties:

(1) $f(0) = 0$ and $f(\frac{1}{2}\pi) = \sqrt{2}$,

(2) $|f(u)-f(v)| \leqslant 2|\sin\frac{1}{2}(u-v)|$ for all u, v in $[0, \frac{1}{2}\pi]$,

(3) $L(f) = \frac{1}{2}\sqrt{2}\,\pi$, where $L(f)$ is the length of the graph of f over $[0, \frac{1}{2}\pi]$.

Since the construction of f is rather complicated, it seems advisable first to sketch briefly the procedure which is followed in the construction.

We start with the function F for which $F(x) = 2\sin\frac{1}{2}x$, $0 \leqslant x \leqslant \frac{1}{2}\pi$. We next define a sequence $g_0, g_1,\ldots$ of 'zigzag' functions on $[0, \frac{1}{2}\pi]$ whose graphs are obtained by piecing together curves which are obtained by translating and possibly reflecting initial portions of the graph of F.

Quart. J. Math. Oxford (2), 4 (1953), 314–20.

The functions g_n are constructed so as to be continuous and have the following properties:

(4) $g_n(0) = 0$ and $g_n(\frac{1}{2}\pi) = \sqrt{2}$ for each n,

(5) $|g_n(u)-g_n(v)| \leqslant 2|\sin\frac{1}{2}(u-v)|$ for each n and all u, v in $[0, \frac{1}{2}\pi]$,

(6) $\lim_{n\to\infty} L(g_n) = \frac{1}{2}\sqrt{2}\,\pi$,

(7) the sequence $g_0, g_1, \ldots$ converges uniformly to a function f,

(8) $\lim_{n\to\infty} L(g_n) = L(f)$.

Once the g_n have been constructed so as to have the above properties, it is easy to see that the function f obtained as the limit of the sequence satisfies conditions (1), (2), (3).

The chief difficulty which arises in our construction is in describing the functions g_n in such a way that one can prove that they satisfy (5).

3. Some preliminary definitions

If g is a function, we let $D(g)$ be the domain of g. If J is a subset of $D(g)$, we define $g|J$ to be the function whose domain is J and which agrees with g on J.

We have already defined the special function F in § 2. If $0 \leqslant a \leqslant \frac{1}{2}\pi$ and b is a real number, we define $P_{a,b}$ to be the function whose domain is $[a, \frac{1}{2}\pi]$ and which satisfies $P_{a,b}(x) = b+F(x-a)$.

We define $N_{a,b}$ to be the function whose domain is $[a, \frac{1}{2}\pi]$ and which satisfies $N_{a,b}(x) = b-F(x-a)$. It is important to note that, if two functions $P_{a,b}$ and $P_{c,d}$ are not equal, then their graphs intersect in at most one point.

We say that a function g is of *type P* on an interval $[a, b] \subset D(g)$ if and only if there exists a number c such that $g|[a, b] = P_{a,c}|[a, b]$. Likewise, a function g is of *type N* on $[a, b]$ if and only if there exists c such that $g|[a, b] = N_{a,c}|[a, b]$.

We define Z to be the set of all continuous functions g such that: the domain of g is a closed interval and is contained in $[0, \frac{1}{2}\pi]$; and $D(g)$ can be subdivided into a finite number of closed intervals on each of which g is either of type P or of type N, the types differing on adjacent intervals of the subdivision. It is easy to see that, if $g \in Z$, then there exists a unique subdivision $\sigma(g)$ of $D(g)$ which has the property that on each interval of $\sigma(g)$ the function is either of type P or of type N.

If $g \in Z$ and $x \in J \in \sigma(g)$, we say that the interval J is a 'g-carrier' of x. Most points of g will have exactly one g-carrier, but a common end-point of two intervals in $\sigma(g)$ will have each of the two intervals as g-carriers.

If g is of type P on $D(g)$ and $h \in Z$, we say that *h is a modification of g* if and only if:

(i) g and h have a common domain $[a, b]$;

(ii) $g(a) = h(a)$ and $g(b) = h(b)$;

(iii) the number of intervals in $\sigma(h)$ is odd and greater than 1;

(iv) if $a = x_0 < x_1 < \dots < x_n = b$ are the end points of the intervals of $\sigma(h)$, then $h(x_k) = g(x_k)$ for all odd k; and

(v) $h(x_k) < h(x_{k+2})$ for all even k.

If g is of type N on $D(g)$ and $h \in Z$, we define h to be a modification of g if and only if $-h$ is a modification of $-g$. If $g \in Z$ and $h \in Z$, we define h to be a modification of g if and only if $h|J$ is a modification of $g|J$ for each interval J in $\sigma(g)$.

We define T to be the set of all ordered triples (g, u, v) for which g is a function, $D(g) \subset [0, \frac{1}{2}\pi]$, and u, v are in $D(g)$. We partially order T by defining $(g, u, v) < (h, s, t)$ if and only if $|u-v| \geqslant |s-t|$ and

$$|g(u)-g(v)| \leqslant |h(s)-h(t)|.$$

We define A to be the set of all (g, u, v) in T for which

$$|g(u)-g(v)| \leqslant 2|\sin \tfrac{1}{2}(u-v)|.$$

It follows at once from the fact that F is increasing on $[0, \frac{1}{2}\pi]$ that if $(g, u, v) < (h, s, t)$ and $(h, s, t) \in A$ then $(g, u, v) \in A$. We shall frequently make use of this fact.

We define B to be the set of all functions g such that $(g, u, v) \in A$ for all u, v in $D(g)$.

Finally, we define inductively a sequence $M_0, M_1, \dots$ of sets of functions. We define M_0 to be the set of all functions g such that $D(g)$ is a closed interval contained in $[0, \frac{1}{2}\pi]$ and g is either of type P or of type N on $D(g)$. If M_n has been defined, then M_{n+1} is the set of all functions g such that g is a modification of some member of M_n.

4. The construction of f

I now prove the crucial lemma.

LEMMA 1. *If n is a non-negative integer, then $M_n \subset B$.*

Proof. I shall prove this lemma by induction on n. For $n = 0$ the lemma is implied by the fact that F is concave and increasing. Before considering the general inductive step, I first consider the special case $n = 1$.

The case $n = 1$. Let $g_1 \in M_1$ and u and v be members of $D(g_1)$. There exists $g_0 \in M_0$ such that g_1 is a modification of g_0, and we may assume

without loss of generality that g_0 is of type P on $D(g_0)$. We also assume $u < v$ and $g_1(u) \neq g_1(v)$.

If u and v have a common g_1-carrier J, then the fact that $g_1|J \in M_0$ implies that $(g_1, u, v) \in A$. We now make the additional assumption that u and v do not have a common g_1-carrier. There are four situations which we must consider.

Suppose that $g_1(u) < g_1(v)$ and u has a g_1-carrier $[a, b]$ on which g_1 is of type P. Then $g_1(v) \leqslant P_{a,c}(v)$ where $c = g_1(a)$. It follows that

$$(g_1, u, v) < (P_{a,c}, u, v),$$

and, since $P_{a,c} \in M_0$, we see that $(g_1, u, v) \in A$.

Suppose $g_1(u) < g_1(v)$ and g_1 is not of type P on any g_1-carrier of u. Then there is an interval $[a, b]$ which is adjacent to and to the right of the (necessarily unique) g_1-carrier of u, and g_1 is of type P on $[a, b]$. There exists $w \in [a, b]$ such that $g_1(w) = g_1(u)$ and $u < w < v$. Thus $(g_1, u, v) < (g_1, w, v)$. It follows from the previous cases which we have considered that $(g_1, w, v) \in A$, and hence $(g_1, u, v) \in A$.

Suppose that $g_1(u) > g_1(v)$ and that u has a g_1-carrier $[a, b]$ on which g_1 is of type N. It follows from condition (*v*) in the definition of modification that $g_1(v) > N_{a,c}(v)$, where $c = g_1(a)$. Thus $(g_1, u, v) < (N_{a,c}, u, v) \in A$ and hence $(g_1, u, v) \in A$.

Suppose that $g_1(u) > g_1(v)$ and that u does not have a g_1-carrier on which g_1 is of type N. Then there exists $[a, b] \in \sigma(g_1)$ which is adjacent to and to the right of the (necessarily unique) g_1-carrier of u, and g_1 is of type N on this interval. Using condition (*v*) together with the facts that $u < v$ and $g_1(u) > g_1(v)$, we see that there must exist $w \in [a, b]$ such that $g_1(w) = g_1(u)$ and $u < w < v$. Thus $(g_1, u, v) < (g_1, w, v)$, and it follows from cases which we have already considered that $(g_1, w, v) \in A$. Hence $(g_1, u, v) \in A$.

The induction step. Now suppose $n > 1$ and assume the lemma for all non-negative integers less than n. Let $g_n \in M_n$ and let u, v be members of $D(g_n)$. We assume $u < v$ and $g_n(u) \neq g_n(v)$. There exist functions $g_0, g_1, \dots, g_n$ such that $g_k \in M_k$ for $k = 0, \dots, n$ and g_{k-1} is a modification of g_k for $k = 1, \dots, n$. There is no loss of generality in assuming that g_0 is of type P on $D(g_0)$.

If for some k $(1 \leqslant k \leqslant n)$ u and v have a common g_k-carrier J, then $g_n|J \in M_{n-k}$, and it follows from the induction assumption that

$$(g_n, u, v) \in A.$$

We now make the additional assumption that u and v do not have a

common g_k-carrier for $1 \leqslant k \leqslant n$. There are six situations which we must consider.

Suppose that $g_n(u) < g_n(v)$; and that for some k $(1 \leqslant k < n)$ u has a g_k-carrier on which g_k is of type N. Then $g_k(u) \leqslant g_n(u)$, and there exists w in some g_k-carrier of v such that $g_k(w) = g_n(v)$, $u < w \leqslant v$. It follows that $(g_n, u, v) < (g_k, u, w) \in A$ and hence $(g_n, u, v) \in A$.

Suppose that $g_n(u) < g_n(v)$; that there is no g_k-carrier of u on which g_k is of type N for $1 \leqslant k < n$; and that u has a g_n-carrier $[a, b]$ on which g_n is of type P. Using the fact that functions $P_{x,y}$ and $P_{s,t}$ can intersect in at most one point, it is easy to see that $g_n(v) \leqslant g_0(v) \leqslant P_{a,c}(v)$, where $c = g_n(a)$. It follows that $(g_n, u, v) < (P_{a,c}, u, v) \in A$ and hence

$$(g_n, u, v) \in A.$$

Suppose that $g_n(u) < g_n(v)$; that there is no g_k-carrier of u on which g_k is of type N for $1 \leqslant k < n$; and that u does not have a g_n-carrier on which g_n is of type P. Then there exists an interval $[a, b] \in \sigma(g_n)$ which is adjacent to and to the right of the (necessarily unique) g_n-carrier of u. We see that g_n is of type P on $[a, b]$ and that there exists w in this interval for which $g_n(w) = g_n(u)$ and $u < w < v$. We obtain $(g_n, u, v) < (g_n, w, v)$, and it follows from cases which we have already considered that

$$(g_n, w, v) \in A.$$

Thus $(g_n, u, v) \in A$.

Suppose that $g_n(u) > g_n(v)$; and that for some k $(1 \leqslant k < n)$ u has a g_k-carrier on which g_k is of type P. Then $g_k(u) \geqslant g_n(u)$. There exists w in some g_k-carrier of v such that $g_k(w) = g_n(v)$ and $u < w \leqslant v$. It follows that $(g_n, u, v) < (g_k, u, w) \in A$ and hence $(g_n, u, v) \in A$.

Suppose that $g_n(u) > g_n(v)$; that there is no g_k-carrier of u on which g_k is of type P for $1 \leqslant k < n$; and that u has a g_n-carrier $[a, b]$ on which g_n is of type N. Then it is not difficult to see that $g_n(v) \geqslant N_{a,c}(v)$, where $c = g_n(a)$. It follows that $(g_n, u, v) < (N_{a,c}, u, v) \in A$ and hence that $(g_n, u, v) \in A$.

Suppose that $g_n(u) > g_n(v)$; that there is no g_k-carrier of u on which g_k is of type P for $1 \leqslant k < n$; and that there is no g_n-carrier of u on which g_n is of type N. There then exists $[a, b] \in \sigma(g_n)$ which is adjacent to and to the right of the (necessarily unique) g_n-carrier of u, and g_n is of type N on $[a, b]$. There exists $w \in [a, b]$ such that $g_n(w) = g_n(u)$ and $u < w < v$. Thus $(g_n, u, v) < (g_n, w, v)$, and it follows from cases which we have already considered that $(g_n, w, v) \in A$. Hence $(g_n, u, v) \in A$.

We have at last completed the proof of Lemma 1.

We now let X be the set of all continuous functions whose domains are $[0, \frac{1}{2}\pi]$. If g and h are members of X we define

$$d(g,h) = \sup_{0 \leqslant t \leqslant \frac{1}{2}\pi} |g(t)-h(t)|.$$

If $g \in X \cap Z$ and h is a modification of g, we say that h is an *ϵ-modification* of g if and only if $d(g,h) < \epsilon$ and the intervals of $\sigma(h)$ all have length less than ϵ. The proof of the following lemma is relatively simple and is omitted.

LEMMA 2. *If $g \in X \cap Z$ and $\epsilon > 0$ then there exists a function h which is an ϵ-modification of g.*

The following lemma expresses the lower semi-continuity of arc length, and is well known.

LEMMA 3. *If $g \in X$, $\epsilon > 0$, and g is rectifiable, then there exists $\delta > 0$ such that, if $h \in X$ and $d(h,g) < \delta$, then $L(h) > L(g)-\epsilon$.*

The final lemma is easy to prove, and its proof is omitted.

LEMMA 4. *If $g_0, g_1, \ldots$ is a sequence of functions in $X \cap B$ and the sequence converges uniformly to a function h, then $h \in X \cap B$.*

We now define inductively a sequence $g_0, g_1, \ldots$ of functions and a sequence $\delta_0, \delta_1, \ldots$ of positive numbers. First we let $g_0 = F$ and $\delta_0 = 1$. If g_n and δ_n have been defined, we choose g_{n+1} to be a $\frac{1}{2}\delta_n$-modification of g_n and then choose a positive number $\delta_{n+1} < \frac{1}{2}\delta_n$ such that, if

$$d(h, g_{n+1}) < \delta_{n+1},$$

then $L(h) > L(g_{n+1})-\delta_n$. The sequence $g_0, g_1, \ldots$ converges uniformly to a continuous function f. Since $g_n \in M_n \subset B$ for each n, it follows from Lemma 4 that $f \in B$. Since δ_n is greater than the lengths of the intervals in $\sigma(g_n)$ for $n > 0$ and $\lim_{n\to\infty} \delta_n = 0$, it is easy to see that $\lim_{n\to\infty} L(g_n) = \frac{1}{2}\sqrt{2}\,\pi$. It is also easy to see that $d(f, g_n) < \delta_n$ for $n > 0$, and hence

$$L(f) > L(g_n)-\delta_{n-1} \quad \text{for } n > 0.$$

Thus $L(f) \geqslant \lim_{n\to\infty} L(g_n)$. The lower semi-continuity of the length function, however, implies that $L(f) \leqslant \lim_{n\to\infty} L(g_n)$. Thus

$$L(f) = \lim_{n\to\infty} L(g_n) = \tfrac{1}{2}\sqrt{2}\,\pi.$$

It is now easy to see that the function f which we have constructed has the properties (1), (2), and (3) of § 2.

We now follow Green's procedure and extend the domain of definition of f so that on the interval $[0, \pi]$ the function is even with respect to $\frac{1}{2}\pi$, and so that on the entire interval $[0, 2\pi]$ the function f has period π.

In cylindrical coordinates, the curve C can now be described by the equations $r = 1$ and $z = f(\theta)$. It is easy to verify that $L(C) = 2\sqrt{2}\,\pi$ and $H(C) = \sqrt{2}$.

REFERENCE

1. J. W. Green, 'Curves encircling a cylinder', *American Math. Monthly*, 60 (1953), 30–1.

ADDENDUM TO

REMARKS ON AHLFORS' DISTORTION THEOREM

By W. K. HAYMAN (*Exeter*)

[*Quart. J. of Math.* (Oxford), 19 (1948), 33–53]

Since this paper was printed I have found out that the sharp bound in Theorem A, one of the main results, was first obtained by O. Teichmüller in his paper: 'Untersuchungen über konforme und quasikonforme Abbildungen', *Deutsche Mathematik* (1938), 621–7.

17 *August* 1953